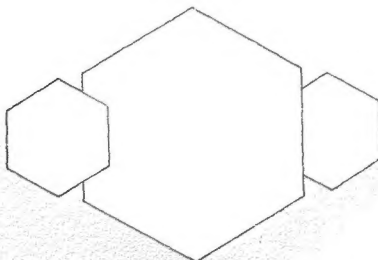
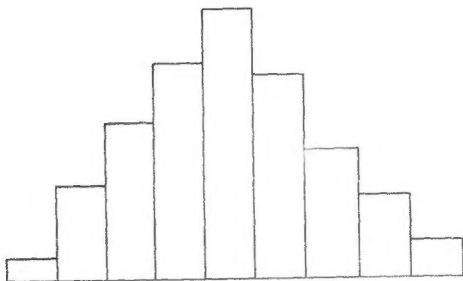


# الاحتمالات والإحصاء

الدكتور محمد خير الدين



مركز النشر العلمي  
جامعة الملك عبد العزيز  
جدة

## اهداءات ٢٠٠٤

المستشار الثقافي السعودي  
محمد عبد العزيز العقيل  
المملكة العربية السعودية





# الْأَخْتِالُ الْإِتِّبَاقُ وَالْإِحْصَاءُ

الشيخ محمد بن عبد العزيز بن باز

قسم الرياضيات الهندسية

كلية الهندسة - جامعة الملك عبد العزيز

مركز النشر العالمي

جامعة الملك عبد العزيز

ص ب ١٥٤٠ - جدة ٢١٢٤١

الهاتف ٥١٢٠٠٠٠ - الفاكس ٥١٢٠٠٠٠

© ١٤١٢ هـ (١٩٩٢ م) جامعة الملك عبد العزيز  
جميع حقوق الطبع محفوظة . غير مسموح بطبع أي جزء من أجزاء  
هذا الكتاب ، أو تخزينه في أي نظام تخزين المعلومات واسترجاعها ، أو نقله  
على أية هيئة أو بآلة وسيلة ، سواء أكانت إلكترونية ، أو شرائط مغنطة ،  
أو ميكانيكية ، أو استنساخاً ، أو تسجيلاً ، أو غيرها إلا بإذن كتابي  
من صاحب حق الطبع .  
الطبعة الأولى : ١٤١٢ هـ (١٩٩٢ م)

مطابع جامعة الملك عبد العزيز

## تقديم

تقوم طريقة البحث في هذا الكتاب على الطريقة المتبعة في البحوث العلمية ، فقد بدأت كاتبي هذا بمقدمة بسيطة عرضت من خلالها الدور الهام الذي لعبته وتلعبه نظرية الاحتمال ، وكذلك فروع هذه النظرية . وضمت كاتبي أيضا ثمانية فصول وذلك تيسيرا للبحث ، وأوردت العديد من التمارين المحلولة وغير المحلولة ، والكثير من الرسوم التي تمكن الطالب من فهم وترسيخ المبادئ والتطبيقات الواردة فيه ، ولقد راعيت في بعض الحالات سرد بعض النظريات ، وذكرت بعض العلاقات دون محاولة لبرهانها ، كذلك راعيت السهولة والبساطة في انتقاء المواضيع الواردة فيه بدقة وبسلسلة لثلاثم ظروف وحاجات طلاب كلية الهندسة والعلوم ، وإمكانيات هذا المقرر .

ولست أدعى بخال أنني استوعبت تماما كل نواحي البحث والدراسة في هذا الموضوع الهام ووصلت فيه إلى الغاية . فمجهود الفرد مهما عظم قليل ، وعلمه مهما خاض ضئيل « ربنا وسعت كل شيء رحمة وعلما » . ولابد من الاعتراف بأن هذا الكتاب جاء وليد مجهودات كثيرة ودراسات استمرت ست سنوات .

لذلك أرجو أن يكون هذا الكتاب فائحة لبحوث جديدة ومنهجيا صالحا لمحاولات صادقة نطمح أن ينهض بها علمائنا .

أسأله تعالى أن يهدينا جميعا إلى صراطه المستقيم وألا يكلنا إلى أنفسنا وأن يوفقنا للوفاء بالعهد والإخلاص في العمل لخير الإنسانية ، إنه نعم المولى ونعم النصير .





## مقدمة

تخضع جميع مظاهر حياتنا للسنن الكونية ، ولعل عجز العلماء عن التنبؤ بنتائج بعض هذه المظاهر من خلال قوانين علمية معروفة دفعهم للقول بأن هذه المظاهر تخضع لعامل الصدفة . فقد ينتهي بنا حادث غير مقصود إلى المستشفى ، ويدفعنا عجزنا عن السيطرة على الصدفة إلى اختيار الطريق الأفضل الذى يليها مباشرة . فنحاول أن نقدر احتمال حدوث حادثة معينة ينمق كل منا بعبارات تدل على الصدفة . فنستعمل كلمات مثل « من المحتمل » ، « ربما » ، أو غيرها ، وعند تفكيرنا بحادثة لم تصبح حقيقة واقعة بعد ، أو كانت نتيجتها خارج نطاق سيطرتنا عليها ، فإننا نقوم عفويا بحساب الاحتمال والصدفة .

لقد نشأت نظرية المصادفة من خلال أبحاث بسكال وفيما المتعلقة بالعباب الحظ المتنوعة . وتبدو فكرة خضوع الصدفة للقوانين غير مقنعة لمن يؤمن بسيطرة ما يسمى بالحظ ، ولكن قوانين الصدفة لا تنفى إمكانية فوز الإنسان في بعض الأحيان بضربة حظ موفقة ، ولا تنكر أيضا قيمة الحُدس في التنبؤ ببعض الحوادث . ولا تصبح هذه القوانين بحق إلا عندما تكثر الأسئلة التى تدخل في البحث ، مثل رميات متعددة لأحجار الترد ، وتوزيع أوراق اللعب عدداً كبيراً من المرات ، واصطدامات كثيرة بين السيارات ، وحساب أعمار عدد كبير من الناس ونشأة الحياة .

لقد استقت نظرية الاحتمال جذورها من المشتغلين بالعباب الحظ الذين كانوا يحاولون استشفاف معلومات تساعدهم على الفوز في لعب الورق والترد . ولم يبدأ حساب الاحتمال بالشكل الذى نعرفه اليوم إلا في منتصف القرن السابع عشر على يد ثلاثة من الفرنسيين هم : دوفرما ، بسكال ، ودوميريه .

وقد برز في هذه الفترة ، لدى الرياضيين عن أبحاثهم في موضوع الاحتمالات ، اختصاص مستقل تماماً عن أعمالهم هو « رياضيات الاحتمال » . وتطور هذا الاختصاص الجديد تطوراً ملموساً حتى أصبح يسيطر على مظاهر عديدة من مظاهر الحياة الحديثة . فإلى جانب سيطرته على التأمين ساعد العالم الذرى في فهم الآثار المتشابكة التي تسجلها الجزيئات الذرية المقذوفة من السيكلوترون على الأفلام ، وساعد خبير الصواريخ في تحديد عوامل الأمان التي يجب أن تزود بها أجهزة القذائف الباهظة الثمن ، وساعد علماء النفس في تقدير ذكاء الأطفال عند إجراء اختبارات الذكاء عليهم ، وساعد رجال الانتخابات في توقع النتائج قبل حدوثها بيوم واحد ، كما ساعد عمال المصانع في تدقيق السلع المنتجة وهى تتدحرج على خطوط الإنتاج في المعمل .

وأصبح نظريات الاحتمال والمصادفة من الأسس الرياضية السليمة ما يجعلها تطبق على نطاق واسع حيثما انعدم الحكم الصحيح المطلق ، وتقدمت دراسة نظريات المصادفة والاحتمال من الوجهة الرياضية تقدماً كبيراً ، حتى أصبحنا قادرين على التنبؤ بحدوث بعض الظواهر التي نقول إنها تحدث بالمصادفة ، والتي لا نستطيع أن نفسر ظهورها بطريقة أخرى ( مثل قذف حجر الترد ) . وأصبحنا بفضل هذه الدراسات قادرين على التمييز بين ما يمكن أن يحدث بطريقة المصادفة وما يستحيل حدوثه بهذه الطريقة .

هذا ويضم الاحتمال بين طياته قانونين هامين ، أولهما يسمى بقانون « أحد الحادئين » ويسمى الآخر بقانون « كلا الحادئين » . كما يضم أيضاً قانوناً هاما يعرف بقانون الأعداد الكبيرة .

وتكمن إحدى الصعوبات الكبيرة عند تطبيق قوانين الاحتمال في تحديد الأساليب الممكنة عليها والتي يمكن أن تحدث بها حادثة ما . وليست هذه المشكلة صعبة جداً في ألعاب الحظ . فقد تمكن علماء الاحتمال النظرى من إيجاد قوانين المتبادلات والمتوافقات والتي سهلت عليهم حل القضايا الصعبة ، بعد تفكيرهم وتأملهم في الأنظمة والتراتب التي يجري بها أى نوع من أنواع اللعب .

وبالرغم من أن حساب الاحتمال لا يزال يحمل آثار اللعب واللهو اللذين اشتق منهما ، إلا أن ألعاب الورق والترد ليست كل شيء فيه . فهو يشكل في مظهره العملي العنصر الرئيس في علم الإحصاء ويدخل حساب الاحتمال الإحصائى مجال الأعمال التجارية ، فيقدر كمية البضائع التى يجب على المنتج أن يخزنها في مخازنه احتياطاً ، ليستطيع تغطية

الطلبات غير متوقعة على متوجاته ، كما يكشف المهندسي المواصلات عدد الاتصالات التي ينبغي أن تنشأ في أى مقسم هاتفى أو شبكة برقية . ويستخدم أيضا في صناعة الأدوية ليكشف عما إذا كانت الآثار التي تنتج من استعمال متطوعين لدواء جديد هي آثار ذات قيمة إحصائية أم أنها مجرد صدفة . ويبقى فرق أساسى بين ألعاب الحظ وبين هذه التطبيقات الأخرى التي تفوق الأولى تعقيدا وتفضلها فائدة . وهو أنه يمكن دائما في ألعاب الحظ أن نعدد كل النتائج الممكنة التي يمكن أن توزع بها أوراق اللعب . وقد يكون هذا التعداد صعبا إلا أنه ممكن دوما . أما عندما نرغب في التنبؤ بتقلبات الحياة ، فكثيرا ما يستحيل أن نعرف مسبقا كل أوراق اللعب المطروحة .

والرياضي الذي يحسب احتمالات الحظ يسحب أوراقه من مجموعة يعرف مسبقا أرباحها ، كما يعرف قيمها النسبية . أما في الاحتمال الإحصائي ، فإن محتويات الجعبة التي يتم السحب منها غير معروفة ، ولابد للإنسان أن ينتقى عينة تجريبية منها وأن يحسن انتقاها ، ثم يقدر احتمال كونها مثلة لمحتويات الجعبة تمثيلا صادقا .

ويشبه حساب الاحتمال ، ومساعدته الإحصاء ، شخصين يتجهان نحو منزل واحد من الناهيتين المختلفتين لطريق واحد . ففي الاحتمال تعرف كل العوامل المؤثرة في القضية . غير أن النتيجة يتكهن بها تكهنا . أما في الإحصاء ، فإن النتائج معروفة ولكن العوامل التي تسببها مشكوك في أمرها . وتوضع إحدى ألعاب حجر النرد حساب الاحتمال إيضاحا حسنا . فحجرا النرد يستطيعان إنتاج 36 توفيقا مختلفة ، ويبقى حساب احتمالاتها طوع يمين كل من يستطيع العد .

يمكن أن نخلص إلى القول بأن علم الاحتمال يبحث في الظواهر العفوية والعلاقات القانونية التي تخضع لها . أما الإحصاء فهو يحوى ضمنا حساب الاحتمال ويحوى إلى جانبه أشياء أخرى تعتبر جزءا مما هو متعارف عليه علميا بالحساب الاحتمالي .

وما برح علم الاحتمال ينمو نموا لا يجهود له ، شأنه في ذلك شأن بقية العلوم ، ولم يسبق أن رمى برج عاجي بظله الطويل على عالم الحياة اليومية ، مثلما فعل برج علم الاحتمال .



# المحتويات

هـ	تقديم .....
ز	مقدمة .....
١	الفصل الأول : الاحتمال .....
٣	فضاء العينة .....
٥	الحوادث .....
٧	العمليات على الحوادث .....
١١	مبادئ العد .....
١٨	احتمال أى حادث .....
٢٢	قوانين الاحتمال .....
٢٥	الاحتمال الشرطي .....
٣٠	نظرية بيز .....
٣٣	تمارين محلولة .....
٤٧	تمارين عامة .....
٥١	الفصل الثاني : المتغيرات العشوائية .....
٥٣	المتغير العشوائى .....
٥٥	التوزيع الاحتمالي المنقطع .....
٦٠	توزيع الاحتمال المستمر .....
٦٤	التوزيعات التجريبية .....
٦٨	توزيع الاحتمال المشترك .....
٨٨	التوقع الرياضي .....
٩٨	قوانين التوقع الرياضي .....

١٠٤	التوقعات الرياضية الخاصة ( التباين - التغير )
١٠٩	خواص التباين
١١٢	نظرية تشييف
١١٥	تمارين محلولة
١٤٨	تمارين عامة
١٥٣	<b>الفصل الثالث : بعض توزيعات الاحتمال المنقطعة</b>
١٥٥	مقدمة
١٥٥	التوزيع المنتظم
١٥٧	التوزيع الحداني والتوزيع المتعدد الحدود
١٦٨	التوزيع الهندسي الزائدي
١٧٧	التوزيع البواسوني
١٨٢	التوزيع الحداني السالب
١٨٦	تمارين محلولة
١٩٦	تمارين عامة
١٩٩	<b>الفصل الرابع : بعض توزيعات الاحتمال المستمرة</b>
٢٠١	التوزيع الطبيعي
٢٠٥	المساحة تحت المنحنى الطبيعي
٢١٤	التقريب الطبيعي للتوزيع الحداني
٢٢١	التوزيعات غما ، الأمي ، كاي مربع
٢٢٧	توزيع وايل
٢٣١	تمارين محلولة
٢٤٢	تمارين عامة
٢٤٥	<b>الفصل الخامس : دوال المتغيرات العشوائية</b>
٢٤٧	تغير المتغيرات
٢٦٠	الدوال المولدة للعزوم
٢٦٨	العينة العشوائية
٢٧١	نظرية العينات

٢٨٠	توزيع المعاينة للوسط
٢٨٦	توزيع المعاينة للمتغير العشوائي $\frac{(n-1)S^2}{\sigma^2}$
٢٨٩	التوزيع t
٢٩٦	التوزيع F
٣٠٢	تمارين محلولة
٣١٣	تمارين عامة
٣١٩	<b>الفصل السادس : نظرية التقدير</b>
٣٢١	مقدمة
٣٢٢	طرق التقدير الكلاسيكية
٣٢٦	تقدير الوسط
٣٣٤	تقدير فرق وسطين
٣٤٣	تقدير P في المجتمع الحداني
٣٤٧	تقدير الفرق بين نسبي مجتمعين حدانيين
٣٥٠	تقدير التباين
٣٥٢	تقدير نسبة تباينين
٣٥٥	تمارين محلولة
٣٦٥	تمارين عامة
٣٦٧	<b>الفصل السابع : اختبارات الفرضيات</b>
٣٦٩	الفرضية الإحصائية
٣٦٩	الأخطاء من النوع الأول I ومن النوع الثاني II
٣٨٣	الاختبارات وحيدة وثنائية الذيل
٣٨٥	اختبار وسط وتباين مجتمع إحصائي
٣٩٥	اختبار حجم العينة لاختبار الوسط
٣٩٨	الاختبارات المتعلقة بالنسب
٤٠٢	اختبار الفرق بين نسبتين
٤٠٥	تمارين محلولة

٤١٧	..... الفصل الثامن : الانحدار والارتباط
٤١٩	..... تمهيد
٤٢٠	..... الانحدار الخطي
٤٢٢	..... الانحدار الخطي البسيط
٤٢٧	..... خواص تقديرات المربعات الصغرى
٤٣٢	..... نهايات الثقة واختبارات المعنوية
٤٣٩	..... تحليل التباين
٤٤١	..... القياسات المتكررة لـ $\gamma$
٤٤٨	..... الارتباط
٤٥٦	..... المراجع
٤٥٧	..... الملاحق
٤٥٩	..... جدول I : المربعات والجنوز التريمية
٤٦٠	..... جدول II : مجموع الاحتمال الحداني $b(x;n, p)$ $\sum_{x=0}^p$
٤٦١	..... جدول III : مجموع الاحتمال البواسوني $p(x;\mu)$ $\sum_{x=0}^{\infty}$
٤٦٤	..... جدول IV : المساحة تحت المنحنى الطبيعي
٤٦٦	..... جدول V : القيم الحرجة في توزيع I
٤٦٧	..... جدول VI : القيم الحرجة في توزيع كاي مربع
٤٦٨	..... جدول VII : القيم الحرجة في توزيع F
٤٧٢	..... جدول VIII : عوامل التحمل في التوزيع الطبيعي
٤٧٥	..... ثبت المصطلحات
٤٧٧	..... عربي / إنجليزي
٤٨٤	..... إنجليزي / عربي



# الفصل الأول

## الاحتمال

■ فضاء العينة ■ الحوادث ■ العمليات على الحوادث ■ مبادئ  
العد ■ احتمال أى حادث ■ قوانين الاحتمال ■ الاحتمال الشرطي  
■ نظرية بيز ■ تقارين محلولة ■ تقارين عامة .



## Sample space العينة (١,١)

يحاول خبراء الإحصاء تفسير نتائج الحظ التي تقابل المشتغلين في الأبحاث العلمية . فهم يهتمون مثلاً بعدد حوادث السير التي تقع شهرياً عند تقاطع شارعي ( فلسطين — المدينة ) في مدينة جدة بالملكة العربية السعودية محاولين تبرير تركيب إشارات ضوئية في هذا التقاطع ، كما أنهم يهتمون بحجم الغاز المنطلق خلال تجربة كيميائية معينة وغير ذلك من الأمور .

إن المعلومات المسجلة بشكلها الخام كأعداد أو قياسات تدعى بالبيانات الخام (raw data) ، فمثلاً تشكيل الأعداد 6, 5, 2, 1, 0, 2 الممثلة لعدد حوادث السير التي وقعت في الثلث الأول من هذا العام عند تقاطع شارعي ( فلسطين — المدينة ) بمدينة جدة مجموعة من البيانات الخام ، كذلك فإن مجموعة القياسات 17, 15, 25 بالسنتيمتر المكعب والمثلة لحجم الغاز المنطلق في عملية كيميائية معينة تمثل أيضاً مجموعة من هذه البيانات .

ويستخدم الإحصائي كلمة تجربة ، أو تجربة إحصائية لوصف أى عملية تقدم له مجموعة بيانات خام . ومن أبسط الأمثلة على التجارب الإحصائيةلقاء قطعة نقود ، إطلاق قذيفة وملاحظة سرعتها في فترات زمنية معينة ، كذلك معرفة الناجحين بالنسبة للمرشح X مثلاً إلخ ..

## تعريف (١,١) فضاء العينة

نسمى مجموعة كل النتائج الممكنة لتجربة إحصائية بفضاء العينة ونرمز لها بالرمز S . إن كل نتيجة في الفضاء S تدعى عنصراً (element) أو نقطة عينة (sample-point) .

إذا احتوى فضاء العينة على عدد محدود من النقاط ، فإن باستطاعتنا ترتيب هذه العناصر ضمن قوسين يفصل بين أى منها فاصلة ، فمثلاً إن فضاء العينة S الممثل لمجموعة نتائج

تجربة إلقاء حجر نرد يمكن كتابته بالشكل :

$$S = \{ 1, 2, 3, 4, 5, 6 \}$$

أما إذا احتوى فضاء العينة على عدد غير محدود من النقاط ، فمن المتعذر عندئذ ذكر جميع عناصره ، وهنا نذكر الخواص التي تحققها عناصر هذا الفضاء ، فمثلا إذا كان  $S$  ممثلا لجميع مدن العالم والتي تعداد سكان أى منها أكثر من نصف مليون نسمة ، فإن  $S$  تكتب على النحو التالي :

$$S = \{ x \mid x \text{ مدينة عدد سكانها أكثر من نصف مليون نسمة} \}$$

كما أن مجموعة النقاط  $(x, y)$  الواقعة على محيط دائرة وداخلها ، نصف قطرها أربعة ومركزها المبدأ يمكن كتابتها على الشكل التالي :

$$S = \{ (x, y) \mid x^2 + y^2 \leq 16 \}$$

إن طريقة تمثيل فضاء العينة بالشكل السابق له منفعة في التجارب التي يكون فيها جدول العناصر ممثلا لعمل شاق روتيني .

### مثال (١,١)

لتكن تجربتنا قذف حجر نرد ، ولنفرض أن اهتمامنا ينصب على العدد الذى يظهر على حجر النرد .  
لدى إلقائه ، فإن :

$$S_1 = \{ 1, 2, 3, 4, 5, 6 \}$$

أما إذا كان اهتمامنا ينصب على نوع العدد زوجى أم فردى ، فعندئذ يكون فضاء العينة من الشكل :

$$S_2 = \{ \text{فردى ، زوجى} \}$$

والمثال (١,١) يوضح لنا أن نتائج تجربة معينة يمكن تمثيلها بأكثر من فضاء عينة واحد / كما أن بعض هذه الأفضية قد تزودنا بمعلومات أكثر مما تزودنا بها بقية الأفضية . فمثلا نلاحظ أن  $S_1$  يحدد لنا جميع عناصر الفضاء ، فإذا علمنا أن عنصراً من  $S_1$  قد وقع فإن

باستطاعتنا أن نحدد أى النتائج في  $S_2$  ستحدث ، ومع ذلك فإن معرفتنا بوقوع حادث في  $S_2$  لن يقدم لنا شيئا عن الحادث الذى وقع في  $S_1$  . وبشكل عام فإننا نستعمل فضاء العينة الذى يقدم لنا أكثر المعلومات حول نتائج التجربة المدروسة .

### مثال (١,٢)

لنسحب أربعة عناصر بشكل عشوائى من مجموعة بضاعة مصنعة . ولنفحص هذه العناصر ، ولتصنف كلا منها بحسب نوعه معابا أو غير معاب . ففى الحالة الأولى نرمز له بالرمز  $D$  ، وفى الثانية بـ  $N$  . نلاحظ أن فضاء العينة الذى يزودنا بمعلومات أكثر هو :

$$S_1 = \{ NNNN, NNND, NNDN, NDNN, DNNN, NNDD, NDDN, NDND, DNND, DNDN, DNNN, NDDD, DNDD, DDND, DDDN, DDDD \}$$

$$S_2 = \{ 0, 1, 2, 3, 4 \} \quad \text{أما الفضاء}$$

فيوضح لنا عدد العناصر المعابة التى سحبناها ، وهو يزودنا بمعلومات أقل من  $S_1$  .

### Events (١,٢) الاحداث

سينصب اهتمامنا فى أى تجربة معطاة حول حدوث حادث معين أكثر من النتيجة الممثلة لعنصر ما فى فضاء العينة ، فمثلا إذا اعتبرنا فى المثال (١,٢) ممثلا على النحو التالى :

$$A = \{ \text{عدد العناصر المعابة أكثر من واحد} \}$$

فهذا يعنى أننا نهم بالمجموعة الجزئية :

$$A = \{ DNNN, DNDN, NNDD, DNND, NDND, NDDN, NDDD, DNDD, DDND, DDDN, DDDD \}$$

من فضاء العينة  $S_1$  . وكذلك الأمر إذا كنا نهم بالحدث  $A$  الذى يقع فيما لو كان الوجه الذى ظهر فى المثال (١,١) يقبل القسمة على 3 فهذا يعنى أيضا أن :

$$A = \{ 3, 6 \}$$

سنخصص بالنسبة لكل حادث مجموعة من نقاط العينة ، وهذه المجموعة هي جزء من فضاء العينة .

### تعريف (١,٢) الحادث The event

الحادث هو مجموعة جزئية من فضاء العينة .

### مثال (١,٣)

بفرض أن  $S = \{t | t > 0\}$  ، حيث يمثل  $t$  العمر السنوي لمادة مشعة ، فإذا كان  $A = \{t | t < 2\}$  فإننا نلاحظ أن  $A$  يمثل مجموعة العناصر المشعة التي تتلاشى قبل نهاية السنة الثانية لتكوينها . هذا المثال يوضح لنا أن أية مجموعة جزئية يمكن أن تحدد لنا حادثاً عناصره نفس عناصر المجموعة الجزئية .

### تعريف (١,٣) الحادث العيني The sample event

إذا حوى حادث ما على نقطة عينة واحدة من فضاء العينة ، قلنا عنه إنه حادث عيني ، أما الحادث المركب فهو الحادث الذى نعبّر عنه من خلال اجتماع عدة حوادث عينية .

### مثال (١,٤)

في تجربة سحب ورقة من ورق اللعب فإن فضاء العينة :

$$S = \{ \text{دينارى ، يستونى ، سباتى ، كبة} \}$$

أما الحادث  $A = \{ \text{دينارى} \}$  فهو يمثل حادثاً عينياً . أما الحادث  $C$  الممثل لسحب ورقة حمراء فهو حادث مركب لأن :

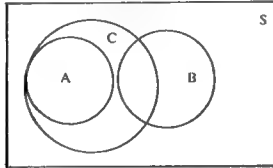
$$C = \{ \text{دينارى ، سباتى} \}$$

### تعريف (١,٤) الفضاء الخالى The empty space or the impossible event

إذا لم يحو فضاء العينة على أية نقطة عينة ، قلنا إنه فضاء صفري أو فضاء خالى ونرمز بالرمز  $\emptyset$

لتكن التجربة سحب كرة سوداء من صندوق يحوى على كرات حمراء فقط . نلاحظ أن فضاء العينة في هذه التجربة هو الفضاء الحالى .

يمكن تمثيل العلاقة بين الحوادث وفضاء العينة بواسطة بما يسمى بمخطط فين ، وفي هذا المخطط تمثل فضاء العينة بمستطيل والمستطيل والحوادث المنبثقة عنه بدوائر في داخله . ففي الشكل (١,١) نلاحظ أن الحوادث A,B,C تمثل جميعها مجموعات جزئية من فضاء العينة S . نلاحظ أيضاً أن الحادث A هو مجموعة جزئية من الحادث C ، كما أن A,B لا يحويان نقاطاً مشتركة ، وأن في الحادثين C,B على الأقل نقطة عينة مشتركة .



الشكل (١,١)

### (١,٣) العمليات على الحوادث Operations on events

يمكن أن نربط بين الحوادث لكي تكون حوادث جديدة ، وذلك باستعمال كافة أنواع العمليات المعروفة على المجموعات .

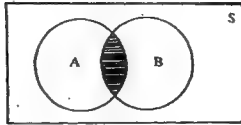
#### تعريف (١,٥) التقاطع The intersection

إن تقاطع حادثين A,B والذي نرمز له بالرمز  $A \cap B$  أو  $A.B$  هو الحادث الذى يقع فيما لو وقع A و B معاً .

والعناصر المنتمية إلى  $A \cap B$  يجب أن تكون موجودة في كل من A و B في نفس الوقت . هذا ويمكن تمثيل التقاطع بواسطة طريقة القاعدة .

$$A \cap B = \{x | x \in A, x \in B\}$$

يوضح القسم المظلل على الشكل (١,٢) عملية التقاطع .



الشكل (١,٢)

مثال (١,٥)

بفرض أن  $A = \{1, 2, 3, 4\}$ ,  $B = \{4, 6, 7\}$  عندئذ نجد أن  $A \cap B = \{4\}$ .

مثال (١,٦)

عند إلقاء قطعة نقود ثلاث مرات وملاحظة نتائج الصورة H والكتابة T فإننا نجد أن :

$$S = \{HHH, HHT, HTH, THH, THT, TTT, TTH, HTT\}$$

بفرض أن A هو الحادث الذى يقع فيما لو ظهر صورتين أو أكثر على التوالي ، وأن B هو الحادث الذى يقع فيما لو ظهرت نفس النتائج فى الرميات الثلاث عندئذ نجد أن :

$$A = \{HHH, HHT, THH\}$$

$$B = \{HHH, TTT\}$$

أما الحادث الممثل لتقاطع الحادثين السابقين فهو الحادث :

$$A \cap B = \{HHH\}$$

وهو الحادث المكون من ظهور الصورة فقط فى الرميات الثلاث . كما نلاحظ أن ظهور الصورة خمس مرات فى هذه التجربة هو المجموعة الخالية أو الحادث المستحيل .

مثال (١,٧)

فى المثال (١,١) وجدنا أن فضاء العينة :

$$S_1 = \{1, 2, 3, 4, 5, 6\}$$

يتكون من ستة نتائج ممكنة للتجربة . فإذا فرضنا أن الحادث A يمثل ظهور وجه زوجي ، B يمثل ظهور وجه فردى ، وأن C يمثل ظهور عدد أولى عندئذ نجد أن :



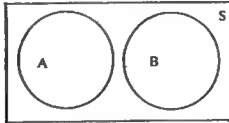
$$A = \{2,4,6\}, B = \{1,3,5\}, C = \{2,3,5\}$$

وأن  $A \cap B = \emptyset$  لأنه لا يمكن ظهور عدد زوجي وفردى بنفس الوقت .

**تعريف (١,٦) الاحداث المتافين تبادلياً** The mutually exclusive events

نقول بأن الحادثين  $A, B$  متافين تبادلياً إذا كان  $A \cap B = \emptyset$  ، وبمعنى آخر إذا تعذر وقوعهما المشترك .

يمكن توضيح مفهوم الاحداث المتافين باستخدام مخطط فين كما في الشكل (١,٣) .



الشكل (١,٣)

نلاحظ في المثال (١,٧) أن الحادثين  $B, A$  متافيان .

**تعريف (١,٧) الاتحاد** The union

إن اتحاد حادثين  $B, A$  والذي نرمز له بالرمز  $A \cup B$  أو  $A + B$  هو الحادث الذي يقع فيما لو وقع  $A$  أو  $B$  أو كلاهما معا .

هذا ويمكن توضيح عناصر الاتحاد بالطريقة التالية :

$$A \cup B = \{x | x \in A \text{ or } x \in B\}$$

يوضح القسم المظلل على الشكل (١,٤) اتحاد الحادثين  $A \cup B$



الشكل (١,٤)

ففي المثال (١,٧) نلاحظ أن  $A \cup B = S$  وأن :

$$A \cup C = \{2,3,4,5,6\}$$

$$B \cup C = \{1,2,3,5\}$$

مثال (١,٨)

بفرض أن :

$$A = \{x | 0 < x < 3\}$$

$$B = \{y | 1 < y < 7\}$$

عندئذ نجد أن :

$$A \cap B = \{z | 1 < z < 3\}$$

$$A \cup B = \{z | 0 < z < 7\}$$

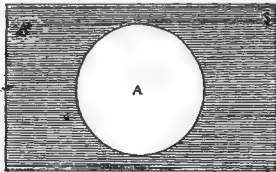
**تعريف (١,٨) متمم حادث ما** The complement of an event

إن متمم أى حادث  $A$  متعلق بالفضاء  $S$  هو مجموعة جميع العناصر من  $S$  غير الموجودة في  $A$ . هذا الحادث الجديد نرمز له بالرمز  $A'$

نلاحظ أن :

$$A' = \{x | x \in S, x \notin A\}$$

يمكن تمثيل المتمم بواسطة مخطط فين كما هو موضح على الشكل (١,٥).



الشكل (١,٥)

نلاحظ في المثال (١,٧) أن متمم الحادث  $A$  هو الحادث  $B$  والعكس صحيح .

من التعاريف السابقة يمكن استنتاج النتائج الهامة التالية :

$$1. A \cup \phi = A$$

$$2. A \cap \phi = \phi$$

$$3. A \cup A' = S$$

$$4. A \cap A' = \phi$$

$$5. (A')' = A$$

$$6. \phi' = S$$

$$7. S' = \phi$$

### (١,٤) مبادئ العد counting principles

سنورد فيما يلي بعض طرق تحديد عدد النواتج الممكنة لتجربة معينة بغير طريقة العد المباشر . وتسمى هذه الطرق باسم التحليل التوافقي .

#### القاعدة الأساسية للعد

##### نظرية (١,١)

إذا أمكن إجراء عملية معينة بعدد من الطرق المختلفة  $n_1$  ، وإذا تلت هذه العملية عملية ثانية وأمکن إجراؤها بعدد  $n_2$  من الطرق المختلفة ، فعندئذ يمكن إجراء العمليتين بعدد من الطرق مساو لـ  $n_1 \cdot n_2$

##### مثال (١,٩)

لدى إلقاء حجرى نرد دفعة ، واحدة ، فإن نقاط فضاء العينة لهذه التجربة هو 36 ، ذلك لأن الحجر الأول يمكن أن يعطى واحدا من ستة نتائج ، ومن أجل كل نتيجة من هذه النتائج ، فإن الحجر الثانى يمكن أن يعطى ستة نتائج أيضا . ولذلك حسب النظرية السابقة ، فإن عدد الثنائيات التى يقدمها الحجران معاً هو 36 . ويوضح الجدول (١,١) هذه النتائج كما يلي :

## جدول (١,١)

النتائج					
(1, 1)	(1, 2)	(1, 3)	(1, 4)	(1, 5)	(1, 6)
(2, 1)	(2, 2)	(2, 3)	(2, 4)	(2, 5)	(2, 6)
(3, 1)	(3, 2)	(3, 3)	(3, 4)	(3, 5)	(3, 6)
(4, 1)	(4, 2)	(4, 3)	(4, 4)	(4, 5)	(4, 6)
(5, 1)	(5, 2)	(5, 3)	(5, 4)	(5, 5)	(5, 6)
(6, 1)	(6, 2)	(6, 3)	(6, 4)	(6, 5)	(6, 6)

هذا ويمكن تعميم النظرية (١,١) لتشمل أى عدد من الحوادث . والنظرية التالية توضح ذلك .

## نظرية (١,٢)

إذا أمكن إجراء عملية ما بعدد  $n_1$  من الطرق المختلفة ، وإذا تلت هذه العملية عملية ثانية وأمكن إجراؤها بعدد  $n_2$  من الطرق المختلفة أيضا ، وإذا تلت هذه العملية الثانية عملية ثالثة وأمكن إجراؤها بعدد  $n_3$  من الطرق المختلفة وهكذا ... عندئذ يكون عدد الطرق التى يمكن أن نجري بها  $k$  عملية دفعة واحدة هو  $n_1 \cdot n_2 \cdot n_3 \dots n_k$  طريقة .

## مثال (١,١٠)

كم عدد فردى مؤلف من ثلاثة أرقام يمكن تشكيله من الأرقام 1,2,5,6,9 إذا استخدم كل رقم مرة واحدة ؟

## الحل

بما أن العدد المطلوب فردى ، فلدينا إذا ثلاثة خيارات من أجل كل رقم آحاد ، ومن أجل كل خيار من هذه الخيارات لدينا أربع خيارات لرقم العشرات وثلاثة خيارات لرقم المئات . وباستخدام النظرية (١,٢) نجد أن عدد الأعداد التى يمكن تشكيلها هو

$$3 \cdot 4 \cdot 3 = 36$$

في أكثر الأحيان ، نهم بفضاءات عينة تحتوي على عناصر من رتب مختلفة أو من فئات مرتبة من العناصر . فمثلا نريد أن نعرف عدد الترتيبات المختلفة الممكنة لجلوس ستة أشخاص على طاولة ، أو نريد أن نسأل السؤال التالي : ما هو عدد التشكيلات الممكنة لسحب ورقتين من مجموع 15 ورقة مرقمة ؟ إن هذه الترتيبات المختلفة تدعى تباديل (permutation)

### تعريف (١,٩) التباديل Permutations

يسمى وضع  $n$  من الأشياء في ترتيب معين بأنه تبديل لهذه الأشياء ( بشرط أن تؤخذ جميع هذه الأشياء دفعة واحدة ) كما يسمى وضع أى عدد  $r \leq n$  من هذه الأشياء في ترتيب معين بأنه تبديل العدد  $n$  من الأشياء مأخوذة  $r$  منها في كل مرة .

فمثلا إن تباديل أربعة حروف مأخوذة جميعها في كل مرة هي :

abcd	bacd	cbad	dabc
abdc	badc	cbda	dacb
acbd	bcad	cdab	dbac
acdb	bcda	cdba	dbca
adbc	bdac	cabd	dcba
adcb	bdca	cadb	dcab

باستخدام النظرية (١,٢) نستطيع أن نصل إلى نفس النتيجة دون تعداد الحالات السابقة . فإذا تصورنا أربعة أمكنة فيمكن ملء المكان الأول بالأحرف الأربعة بأربع طرق ، أما المكان الثاني فيمكن ملؤه بثلاثة طرق ، والثالث بطريقتين ، والرابع بطريقة واحدة وهكذا يمكننا أن نملأ الأمكنة الأربعة دفعة واحدة بعدد من الطرق مساو لـ (4)  $24 = (1) (2) (3)$  تبديلة . وبشكل عام فإن  $n$  من العناصر المختلفة يمكن ترتيبها بـ  $1 \cdot 2 \cdot 3 \cdot \dots \cdot (n-2) \cdot (n-1) \cdot n$  طريقة ، سنرمز للمضروب السابق بالرمز  $n!$  ونقرؤه  $n$  عاملي . فمثلا يمكن ترتيب أربعة عناصر بـ  $4! = 4 \cdot 3 \cdot 2 \cdot 1$  طريقة . من التعريف السابق نجد أن  $1! = 1$

## نظرية (١,٣)

إن عدد تبديل  $n$  من الأشياء المختلفة هو  $n!$  . لنرى مثلاً عدد التباديل الممكنة لحرفين من أربعة حروف  $a, b, c, d$  إن هذه التباديل هي :

$$ab, ac, ad, ba, ca, da, bc, cb, bd, db, cd, dc$$

ولشرح هذه الفكرة نتصور مكانين أو صندوقين . أما المكان الأول فيمكن ملؤه بأربعة حروف ، وبعد ملء المكان الأول بأحد الحروف الأربعة المفروضة ، فإنه يمكن ملء المكان الثاني بثلاثة طرق ، وحسب النظرية (١,٢) نجد أن عدد طرق ملء المكانين السابقين هو  $4 \cdot 3 = 12$  تبديلة . وبشكل عام نجد أن عدد تبديلات  $n$  شيئاً مأخوذاً منها  $r$  شيئاً في وقت واحد هو  $(n - r + 1) \dots (n - 2) (n - 1) n$  طريقة .

سنرمز لهذا المضروب بالرمز

$${}_n P_r = \frac{n!}{(n - r)!}$$

## نظرية (١,٤)

إن عدد تبديل  $n$  شيئاً مختلفة مأخوذة  $r$  شيئاً منها في نفس الوقت هو :

$${}_n P_r = \frac{n!}{(n - r)!}$$

## مثال (١,١٢)

ما هو عدد التباديل المكونة من ستة أشياء مأخوذة ثلاثة ثلاثة ؟

## الحل

إن عدد التباديل هذه هو :

$${}_6 P_3 = \frac{6!}{(6 - 3)!} = 120$$

وتعليل ذلك أنه لو تصورنا ثلاثة أمكنة فإنه يمكن شغل المكان الأول بستة طرق بوساطة هذه الأشياء الستة ، وبعد شغل المكان الأول بأحد هذه الأشياء ، يمكن شغل المكان الثاني بإحدى خمس طرق ، وبعد شغل المكانين الأول والثاني يمكن شغل المكان الأخير بأربع طرق وبحسب النظرية (١,٢) نجد أن هناك  $6 \cdot 5 \cdot 4 = 120$  طريقة .

### التباديل مع التكرار Permutations with repetition

يطلب في بعض الأحيان معرفة عدد تباديل مجموعة من العناصر يكون بعضها متماثلاً فمثلاً وجدنا أن عدد تباديل أربعة حروف  $a, b, c, d$  مختلفة مأخوذة جميعها دفعة واحدة هو 24 تبديلة مختلفة . فإذا فرضنا أن  $a = b, c = d$  عندئذ يمكننا فقط ذكر التباديل التالية :

aacc, acac, caac, ccaa, acca, caca

وهذا العدد من التباديل المختلفة هو  $\frac{4!}{2!2!} = 6$  تبديلة مختلفة .

### نظرية (١,٥)

إن عدد التباديل المختلفة لـ  $n$  شيئا مأخوذة من  $k$  صنفاً بحيث تحتوي  $n_1$  عنصراً من عناصر الصنف الأول ،  $n_2$  من الثاني و .....  $n_k$  من الصنف الأخير هو

$$\frac{n!}{n_1! n_2! \dots n_k!}$$

### مثال (١,١٢)

ما هو عدد الإشارات المختلفة التي يمكن تكوينها من بين ستة أعلام مرفوعة رأسياً بحيث يكون أربعة منها خضر واثنان بيض ؟

### الحل

نلاحظ أنه توجد 15 إشارة مختلفة من أربعة أعلام خضر واثنان بيض .

$$\frac{6!}{4!2!} = 15$$

### مثال (١,١٣)

ما هو عدد التباديل المختلفة التي يمكن تكوينها من جميع أحرف كلمة probability :  
نلاحظ أنه يوجد إحدى عشر حرفاً منها اثنان متشابهان هما  $b$  واثنان آخران متشابهان هما  $i$  وعدد التباديل المختلفة هو :

$$\frac{11!}{2!2!} = 9979200 \text{ تبديلة .}$$

نهتم عادة بعدد الطرق التي يمكن فيها تقسيم  $n$  شيئا إلى  $r$  مجموعة جزئية ( تدعى كل واحدة منها خلية ) وذلك باعتبار أن ترتيب هذه العناصر في كل خلية غير هام . فمثلا إذا كان لدينا مجموعة الأحرف  $\{x_1, x_2, x_3, x_4\}$  فإننا نلاحظ أن عدد الطرق التي يمكن بها تجزئة هذه المجموعة إلى خليتين تحتوي الأولى منها على ثلاثة عناصر ، والثانية على عنصر واحد هو أربع طرق ، وهذه الطرق هي :

$$\{(x_1, x_2, x_3), x_4\}, \{(x_1, x_2, x_4), x_3\},$$

$$\{(x_1, x_3, x_4), x_2\}, \{(x_2, x_3, x_4), x_1\},$$

نرمز للعدد السابق بالرمز  $\binom{4}{3,1} = \frac{4!}{3!1!} = 4$  حيث يمثل العدد العلوي عدد العناصر أما العددين السفليين فيمثلا على الترتيب عدد عناصر الخليتين الأولى والثانية .

### نظرية (١,٦)

إن عدد الطرق التي يتم بها تجزئة مجموعة مؤلفة من  $n$  شيئا مختلفا إلى  $r$  خلية ، تحتوي الأولى على  $n_1$  شيئا ، الثانية  $n_2$  شيئا ، وتحتوي الأخيرة على  $n_r$  شيئا هو

$$\binom{n}{n_1, n_2, \dots, n_r} = \frac{n!}{n_1! n_2! \dots n_r!}$$

حيث إن  $n_1 + n_2 + \dots + n_r = n$

### مثال (١,١٤)

ما هو عدد الطرق التي يمكن بها توزيع 12 قطعة نقدية مختلفة على ثلاثة أطفال ، بحيث يأخذ الطفل الصغير أربع قطع والأوسط خمس والكبير ثلاث قطع ؟

### الحل

نلاحظ أن عدد التجزئات المرتبة لاثنتي عشرة قطعة نقدية إلى ثلاث خلايا تحتوي على الترتيب 3,5,4 هو



$$\frac{12!}{4!5!3!} = 27750$$

في كثير من المسائل العملية نهم بعدد الطرق التي يمكن أن نسحب بواسطتها  $r$  عنصرا من  $n$  عنصرا مفروضا دون الاهتمام بالترتيب . هذه السحبات تدعى بالتوافيق (combinations) ، والتوافق هي بالضبط تجزئة مجموعة مؤلفة من  $n$  عنصرا مختلفا إلى خليتين تحتوي الأولى على  $r$  شيئا ، كما تحتوي الثانية على  $(n-r)$  شيئا المتبقية . سنرمز لهذا العدد من التوافيق بالرمز  $\binom{n}{r}$

### نظرية (١,٧)

إن عدد توافيق  $n$  شيئا متميزا مأخوذة منها  $r$  شيئا في وقت واحد هو

$$\binom{n}{r} = \frac{n!}{r!(n-r)!}$$

### مثال (١,١٥)

رجل له ثمانية عشر ولداً . ما هو عدد الطرق التي يمكنه فيها أن يدعو خمسة من أولاده إلى رفقته في رحلة سيقوم بها إلى بلد أجنبي ؟

### الحل

نلاحظ أن عدد الطرق التي يمكن أن يختار بها خمسة من 18 عنصرا مختلفا هو العدد

$$\binom{18}{5} = \frac{18!}{5!(18-5)!} = \frac{18!}{5! \cdot 13!} = 8568$$

### مثال (١,١٦)

كم لجنة من ثلاثة أشخاص يمكن تشكيلها من خمسة أطفال وثلاث بنات بدون تحديد ؟

### الحل

إن عدد اللجان هو

$$\binom{8}{3} = \frac{8!}{3!5!} = 56$$

### (١,٥) احتمال أى حادث Probability of an event

الاحتمال هو أحد المفاهيم الأساسية الهامة في الوقت الحاضر . نفرض أن كيساً يحتوي على ست كرات مختلفة الألوان متجانسة ومتماثلة ، ثلاث منها حمراء واثنان زرق ، وواحدة بيضاء ، مخلوطة خلطاً جيداً داخل الكيس . إن إمكانية سحب كرة ملونة بصورة عشوائية ( أى سحب كرة حمراء أو زرقاء ) هو أكبر من إمكانية سحب كرة بيضاء . والسؤال المطروح الآن : هل يمكننا وصف هذه الإمكانية بعدد ؟ من الواضح أن هذا ممكناً . إن العدد الذى يصف لنا هذه الإمكانية يدعى احتمالاً (probability) .

لنفرض أن فضاء العينة  $S$  هو فراغ منته ، سنفرق بكل نقطة عينة  $x_i$  من هذا الفضاء عدداً نرمز له بالرمز  $w_i$  بحيث يكون  $\sum w_i = 1$  . فمثلاً في تجربة إلقاء حجر نرد نلاحظ أن جميع أوجه الحجر لها نفس إمكانية الوقوع ولذلك ، فإننا نفرق بكل وجه عدداً يساوى  $\frac{1}{6}$  . نسميه احتمال ظهور ذلك الوجه ، أما في تجربة إلقاء قطعة نقود ، فإننا نلاحظ أن ظهور الصورة (H) له نفس إمكانية ظهور وجه الكتابة (T) ، ولذلك فإننا نفرق بكل وجه عدداً مساوياً لـ  $\frac{1}{2}$  ، يمثل احتمال ظهور كل وجه . إن هذا النوع من فضاءات العينة يدعى بالفضاء الذى الاحتمالات المتساوية أو الفضاء المنتظم . وعلى وجه الخصوص إذا احتوى الفضاء المنتظم  $S$  على  $n$  من النقاط عندئذ نفرق بكل نقطة عينة من هذا الفضاء عدداً  $\frac{1}{n}$  يكون ممثلاً لاحتمال ظهور هذه النقطة .

لإيجاد احتمال وقوع حادث  $A$  نحسب مجموع كافة الأعداد المرفقة بنقاط هذا الحادث . نسمى هذا العدد بقياس الحادث  $A$  أو احتمال الحادث  $A$  ونرمز له بالرمز  $P(A)$  . وهكذا نجد أن قياس  $P$  هو الصفر وقياس  $S$  هو الواحد .

### تعريف (١,١٠) احتمال أى حادث Probability of an event

احتمال أى حادث  $A$  هو مجموع الأعداد المرفقة بمختلف نقاط العينة الموجودة في هذا الحادث .

من هذا التعريف نستنتج النتائج الهامة التالية :

1.  $P(\phi) = 0$
2.  $P(S) = 1$
3.  $0 \leq P(A) \leq 1$

مثال (١٧، ١)

ألقيت قطعة نقود متوازنة ومتماثلة ثلاثة مرات متتالية . ما هو احتمال ظهور الصورة مرة واحدة على الأقل ؟

الحل

نلاحظ أن فضاء العينة S في هذه التجربة هو :

$$S = \{ HHH, HTH, THH, HHT, TTH, THT, HTT, TTT \}$$

وبما أن القطعة متوازنة ، إذاً كل نتيجة ممكنة من هذه النتائج لها نفس إمكانية الوقوع .  
لذلك فإننا نفرق كل نقطة عينة بعدد w . وهكذا نجد أن  $8w = 1$  أي إن  $w = \frac{1}{8}$   
لنرمز للحدث المطلوب بالرمز A ، عندئذ يكون :

$$P(A) = \frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8} = \frac{7}{8}$$

مثال (١٨، ١)

صمم حجر نرد بحيث أن ظهور أى وجه زوجي له ضعف إمكانية ظهور أى وجه فردى . ما هو احتمال ظهور وجه أقل من 5 عند إلقاء هذا الحجر مرة واحدة ؟

الحل

$$S = \{ 1, 2, 3, 4, 5, 6 \}$$

لنفرق بكل عدد فردى العدد w وبكل عدد زوجي من S العدد 2w بحيث يكون  $9w = 1$   
نلاحظ أن  $w = \frac{1}{9}$  ، والاحتمال المطلوب للحدث A الممثل لظهور وجه أقل من 5 هو

$$P(A) = \frac{1}{9} + \frac{2}{9} + \frac{1}{9} + \frac{2}{9} = \frac{6}{9} = \frac{2}{3}$$

يمكن النظر إلى الأعداد المرفقة بنقاط أى حادث  $A$  على أساس أنها تمثل احتمالات الحوادث العينية التي يتألف منها الحادث  $A$  ، فإذا كانت التجربة من النوع الذى يمكن أن نفرض فيها أوزاناً متساوية لكل نقطة عينة من  $S$  ، فعندئذ يكون احتمال أى حادث وليكن  $A$  ممثلاً للنسبة بين عدد عناصر الحادث  $A$  وعدد عناصر فضاء العينة  $S$  .

### نظرية (١,٨)

إذا كان عدد إمكانيات تجربة معينة  $N$  إمكانية ( متساوية فى إمكانية وقوعها ) منها  $n$  إمكانية توافق الحادث  $A$  فعندئذ يكون احتمال هذا الحادث مساوياً .

$$P(A) = \frac{n}{N} : n \leq N$$

### مثال (١,٩)

أهو احتمال الحصول على ورقة ديتارى عند سحب ورقة من ورق اللعب بصورة عشوائية ؟

### الحل

إن عدد إمكانيات سحب ورقة من ورق اللعب هو  $N = 52$  .

ويوافق الحادث  $A$  المثل لظهور ورقة ديتارى عدداً من الحالات  $n = 13$  ، وعليه فإن احتمال  $A$  هو

$$P(A) = \frac{n}{N} = \frac{13}{52} = \frac{1}{4}$$

### مثال (١,٢٠)

ما هو احتمال ظهور وجه الصورة عند إلقاء قطعة نقود متوازنة ومتماثلة ؟

### الحل

نلاحظ أن عدد النتائج الممكنة لدى إلقاء قطعة نقود متوازنة هو  $N = 2$  ويوافق الصورة

من هذه النتائج  $n = 1$  نتيجة ، وهكذا نجد أن احتمال الحادث B الممثل لظهور وجه الصورة هو

$$P(B) = \frac{n}{N} = \frac{1}{2}$$

مثال (١,٢١)

ما هو احتمال حصولنا على مجموع يساوى خمسة عند إلقاء حجرى نرد مرة واحدة ؟

الحل

بالعودة إلى المثال (١,١٠) نجد أن فضاء العينة يتألف من 36 نقطة عينة ويوافق الحادث A الممثل لظهور مجموع يساوى خمسة في الثنائيات التالية :

$$(2,3), (3,2), (4,1), (1,4)$$

وعدها أربع نتائج ممكنة . وهكذا نجد أن احتمال الحادث هو

$$P(A) = \frac{4}{36} = \frac{1}{9}$$

مثال (١,٢٢)

اخترنا نقطة بطريقة عشوائية من داخل دائرة . ما هو احتمال أن تكون هذه النقطة أقرب إلى مركز الدائرة منها إلى المحيط ؟

الحل

نفرض أن S مجموعة النقاط الواقعة داخل الدائرة ذات نصف القطر r ، وأن A مجموعة النقاط الواقعة داخل الدائرة المشتركة مع الدائرة الأولى في نفس المركز ونصف قطرها  $\frac{1}{3}r$  عندئذ

$$P(A) = \frac{\text{مساحة } A}{\text{مساحة } S} = \frac{M\left(\frac{1}{3}r\right)^2}{M r^2} = \frac{1}{9}$$

### (١,٦) قوانين الاحتمال Probability Formulas

يبدو في بعض الأحيان أن من السهل حساب احتمال حادث من الاحتمالات المعروفة لبعض الحوادث الأخرى . وهذا يبدو جليا إذا كان الحادث المراد حساب احتماله اجتماعا لحوادث أخرى معلومة الاحتمالات ، أو إذا كان متمما لحادث علم احتماله . سنذكر الآن بعض القوانين التي توفر لنا على الغالب حساب الاحتمال ، وأول هذه القوانين يدعى بقانون الجمع .

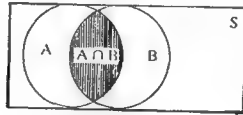
#### نظرية (١,٩)

لأى حادثين  $A, B$  فإن

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

#### البرهان

بالعودة إلى مخطط فين الموضح بالشكل (١,٦) نجد أن  $P(A \cup B)$  هو مجموع الأعداد المرفقة بنقاط العينة للحادث  $A \cup B$  . غير أن  $P(A) + P(B)$  يمثل مجموع كل الأعداد المرفقة بنقاط العينة الموجودة في الحادث  $A$  زائدا مجموع كل الأعداد المرفقة بنقاط العينة الموجودة في  $B$  . لذلك فإننا نجمع الأعداد المرفقة بنقاط العينة الموجودة في  $A \cap B$  مرتين . وهذه الأعداد المرفقة بالتقاطع يجب طرحها من  $P(A) + P(B)$  وهي تساوى  $P(A \cap B)$  وبذلك نحصل على المطلوب .



الشكل (١,٦)

#### نتيجة (١,١)

إذا كان الحادثان  $A, B$  متنافيين فعندئذ يكون :

$$P(A \cup B) = P(A) + P(B)$$

إن النتيجة (١,١) تتج مباشرة من النظرية (١,١٠) لأنه إذا كان  $B, A$  متنافيين فعندئذ يكون  $A \cap B = \phi$  ، ومنه  $P(A \cap B) = 0$  . وبشكل عام يمكننا أن نكتب النتيجة التالية :

#### نتيجة (١,٢)

إذا كانت الحوادث  $A_1, A_2, \dots, A_n$  متنافية فعندئذ يكون :

$$P(A_1 \cup A_2 \cup \dots \cup A_n) = P(A_1) + P(A_2) + \dots + P(A_n)$$

لاحظ أنه إذا كانت جملة الحوادث  $A_1, A_2, \dots, A_n$  مؤلفة لتجزئة لفضاء العينة  $S$  ، فعندئذ يكون :

$$P(A_1 \cup \dots \cup A_n) = P(A_1) + \dots + P(A_n) = P(S) = 1$$

#### مثال (١,٢٣)

بفرض احتمال أن يجتاز طالب امتحان الرياضيات هو 0.7 ، و امتحان الفيزياء هو 0.8 . فإذا علمت احتمال أن يجتاز واحد على الأقل من الامتحانين هو 0.9 ، فما هو احتمال أن يجتاز كلا الامتحانين ؟

#### الحل

إذا كان  $A$  ممثلاً لاجتياز الطالب امتحان الرياضيات و  $B$  امتحان الفيزياء ، عندئذ نجد من النظرية (١,١٠) أن :

$$\begin{aligned} P(A \cup B) &= P(A) + P(B) - P(A \cap B) \\ &= 0.7 + 0.8 - 0.9 = 0.6 \end{aligned}$$

#### مثال (١,٢٤)

احسب احتمال الحصول على مجموع يساوى 5 أو 10 لدى إلقاء زوج من أحجار النرد .

#### الحل

لنفرض الحادث  $A$  الممثل لمجموع يساوى 5 ، والحادث  $B$  لمجموع يساوى 10 . نلاحظ

أنه يوافق الحادث A أربع نقاط عينية ( من فضاء العينة المكون من 36 نقطة ) انظر المثال (١,١٠) وهذه النقاط هي

$$(1,4),(4,1),(3,2),(2,3)$$

كما يوافق الحادث B ثلاث نقاط هي (4,6),(6,4), (5,5) ، ونظراً لكون هذه النقاط جميعها متساوية الإمكانية فإننا نجد أن  $P(A) = \frac{1}{9}$   $P(B) = \frac{1}{12}$  . كما نلاحظ أن الحادثين B,A متنافيان (لأنه لا يمكن الحصول على مجموع 10,5 بنفس الوقت في نفس الإلقاء ) لذلك نجد أن

$$P(A \cup B) = P(A) + P(B)$$

$$= \frac{1}{9} + \frac{1}{12}$$

$$= \frac{7}{36}$$

نظرية (١,١٠)

بما أن A, A' يمثلان حادثين متنافيين عندئذ يكون :

$$P(A') = 1 - P(A)$$

البرهان

نلاحظ أن  $A \cup A' = S$  إذا

$$1 = P(S)$$

$$= P(A \cup A') = P(A) + P(A')$$

لذلك فإن :

$$P(A') = 1 - P(A)$$

مثال (١,٢٥)

ألقيت قطعة نقود متأللة ومتوازنة أربع مرات متتالية . ما هو احتمال الحصول على وجه الصورة مرة على الأقل ؟



## الحل

نفرض أن الحادث  $A$  يمثل ظهور وجه الصورة مرة على الأقل . نلاحظ أن فضاء العينة  $S$  يحتوى على عدد من نقاط العينة  $2^4 = 16$  نقطة لانه يقابل كل إلقاء إحدى نتيجتين ، كما نلاحظ أن  $P(A) = 1 - P(\bar{A})$  حيث أن  $\bar{A}$  يمثل عدم ظهور أى صورة . ويمكن أن يقع  $\bar{A}$  بشكل واحد فقط ولذلك فإن :

$$P(\bar{A}) = \frac{1}{16}$$

$$P(A) = 1 - \frac{1}{16} = \frac{15}{16} \quad \text{وأخيراً فإن}$$

## Conditional probability (١,٧) الاحتمال الشرطى

إن احتمال وقوع حادث  $B$  عند معرفتنا بأن حادثاً  $A$  قد وقع يدعى بالاحتمال الشرطى ونرمز له بالرمز  $P(B|A)$  ، ونقرأ احتمال وقوع  $B$  علماً أن الحادث  $A$  قد وقع ، أو بشكل أبسط احتمال  $B$  بمعلومية  $A$  .

لنفرض أن  $B$  يمثل ظهور عدد فردى عند إلقاء حجر نرد مرة واحدة . معتبرين أن حجر النرد مصمم بحيث تكون امكانية ظهور وجه فردى مساوية لضعف إمكانية ظهور وجه زوجى . فإذا أرفقنا بكل وجه زوجى العدد  $\frac{1}{9}$  وبكل وجه فردى العدد  $\frac{2}{9}$  فإننا نجد أن احتمال الحادث  $B$  هو  $P(B) = \frac{2}{3}$  ، لنفرض الآن أننا نعلم أن نتيجة الإلقاء هى عدد أكبر من 2 . إذا نحن أمام فضاء عينة جديدة  $A = \{3,4,5,6\}$  . فلإنيجاد احتمال وقوع الحادث  $B$  بالنسبة للفضاء الجديد  $A$  ، علينا أن نرفق أعداداً جديدة بعناصر هذا الفضاء بحيث يكون مجموع هذه الأعداد الواحد . فإذا أرفقنا بأى عدد زوجى  $w$  ، وبأى عدد فردى  $2w$  ، فإننا نجد أن  $6w = 1$  ، ومنه  $w = 1/6$  . وبالنسبة للفضاء  $A$  ، فإننا نجد أن  $B$  يحوى العنصرين 3,5 .

$$P(B/A) = \frac{4}{6} = \frac{2}{3}$$

هذا المثال يوضح أن لنفس الحادث  $B$  احتمالات مختلفة بالنسبة لفضاءات مختلفة . ويمكننا أيضاً أن نكتب :

$$P(B/A) = \frac{2}{3} = \frac{4/9}{6/9} = \frac{P(A \cap B)}{P(A)}$$

حيث حسبنا  $P(A)$ ,  $P(A \cap B)$  من فضاء العينة الأساس  $S$ . وبعبارة أخرى فإن الاحتمال الشرطي بالنسبة للفضاء الجزئي  $A$  من  $S$  يمكن حسابه مباشرة من الفضاء  $S$  نفسه.

### تعريف (١,١١)

نعرف الاحتمال الشرطي للحدث  $B$  بمعلومية  $A$ ، والذي نرمز له بالرمز  $P(B/A)$

بالعلاقة :

$$P(B/A) = \frac{P(A \cap B)}{P(A)}$$

شرطية أن يكون

$$P(A) > 0$$

### مثال (١,٢٦)

ما هو احتمال الحصول على ثلاث صور لدى إلقاء ثلاث قطع نقود إذا علمت أنه قد ظهر صورة على القطعة الأولى لدى إلقائها ؟

### الحل

نلاحظ أن فضاء العينة  $S$  يتألف من ثمانى نقط هي

$$S = \{HHH, TTT, HTT, THT, TTH, HHT, HTH, THH\}$$

فإذا كان على القطعة الأولى صورة ، فإننا نجد أن فضاء العينة الجزئي  $A$  هو

$$A = \{HHH, HTT, HHT, HTH\}$$

وحيث أن الحادث  $B$  الممثل لظهور ثلاث صور يولفقه حالة واحدة في الفضاء  $A$  الذى يحتوى على أربع نقاط فإنا نجد أن

$$P(B/A) = \frac{P(A \cap B)}{P(A)} = \frac{1/8}{4/8} = \frac{1}{4}$$

الرمز  $P(B/A)$  يمثل احتمال ظهور ثلاث صور بمعلومية أن القطعة الأولى قد أظهرت صورة .

## مثال (١,٢٧)

عند إلقاء حجرى نرد متماثلين ومتوازنين ، وجد أن الوجهين الظاهرين مختلفان . ماهو احتمال :

- (١) أن يكون مجموع الوجهين اللذين ظهرهما خمسة ؟
- (٢) أن يكون مجموع الوجهين أقل أو يساوى خمسة ؟

## الحل

بالعودة إلى مثال (١,١٠) نلاحظ أن فضاء العينة S يحتوى على 36 نقطة عينة بينها ست نقاط يكون فيها الوجهان متماثلين وهذه النقاط هي :

$$(1,1),(2,2),(3,3),(4,4),(5,5),(6,6)$$

ولهذا فإن فضاء العينة المختزل A يحتوى على  $30 = 36 - 6$  نقطة عينة ، فإذا فرضنا أن الحادث B يمثل مجموع الوجهين الظاهرين خمسة فإننا نجد أنه يوافق هذا الحادث في الفضاء المختزل أربع نقاط هي :

$$(2,3),(3,2),(4,1),(1,4)$$

وهكذا نجد أن الاحتمال الأول :

$$P(B) = \frac{4}{30} = 0.1333$$

أما إذا رمزنا للحادث الممثل لظهور مجموع يساوى خمسة أو أقل بالرمز C ، فإننا نجد أنه يوافق الحادث C النقاط التالية :

$$(1,2),(2,1),(1,3),(3,1),(1,4),(4,1),(2,3),(3,2)$$

ومنه :

$$P(C) = \frac{8}{30} = 0.2666$$

لاحظ أن الاحتمالات المحسوبة هي احتمالات شرطية بالنسبة لفضاء العينة غير المختزل . إذا ضربنا العلاقة الواردة في التعريف (١,١١) بالاحتمال P(A) ، فإننا نستنتج نظرية الضرب التالية :

## نظرية (١,١١)

إذا أمكن وقوع كلا الحادتين B,A في تجربة معينة ، عندئذ نجد أن احتمال وقوعهما المشترك يعطى بالعلاقة التالية :

$$P(A \cap B) = P(A) \cdot P(B/A)$$

ولتوضيح طريقة استخدام النظرية (١,١١) نفرض أن لدينا علبة تحوى 15 فيوزا وبين هذا العدد من الفيوزات ثلاثة عاطلة ( معابة ) . لنسحب فيوزين بصورة عشوائية من العلبة على التالى وبدون إعادة الفيوز الأول المسحوب إلى العلبة . ما هو احتمال أن يكون الفيوزان المسحوبان عاطلين ( معابين ) ؟ وللإجابة على هذا السؤال سنعتبر أن الحادث A يمثل أن الفيوز الأول المسحوب عاطل ، وأن B : يمثل الفيوز الثانى المسحوب عاطل أيضا . نلاحظ أن احتمال A هو  $\frac{1}{5}$  ، وأن  $P(B/A) = \frac{1}{7}$  ولذلك فإن :

$$P(A \cap B) = P(A) \cdot P(B/A) = \frac{1}{5} \cdot \frac{1}{7} = 0.0526$$

ولتعميم النظرية (١,١١) نسوق النظرية التالية :

## نظرية (١,١٢)

إذا أمكن في تجربة معينة وقوع كل الحوادث  $A_1, A_2, \dots, A_n$  دفعة واحدة فعندئذ يكون احتمال الوقوع المشترك لهذه الحوادث جميعها معطى بالعلاقة التالية :

$$P(A_1 \cap A_2 \cap \dots \cap A_n) = P(A_1) P(A_2/A_1) P(A_3/A_1 \cap A_2) \dots$$

$$\dots P(A_n/A_1 \cap A_2 \cap \dots \cap A_{n-1})$$

## ملاحظة

في المثال السابق إذا أعيد الفيوز الأول المسحوب إلى العلبة فإننا نجد أن :

$$P(B/A) = P(B)$$

وهذا يعنى أن الحادثين B,A مستقلان .

**تعريف (١,١٢) استقلال الحادثين Independence of events**

إن الشرط اللازم والكافى ليكون الحادثان B,A مستقلين هو أن يتحقق ما يلى :

$$P(A \cap B) = P(A) \cdot P(B)$$

**مثال (١,٢٨)**

لنحسب احتمال الحصول على مجموع يساوى 5 و 10 وذلك عند إلقاء حجرى نرد متوازنين ومتماثلين مرتين .

**الحل**

لنرمز للحوادث التالية :

$A_1$  = { ظهور مجموع يساوى 5 على الوجهين فى القذفة الأولى }

$A_2$  = { ظهور مجموع يساوى 5 على الوجهين فى القذفة الثانية }

$B_1$  = { ظهور مجموع يساوى 10 على الوجهين فى القذفة الأولى }

$B_2$  = { ظهور مجموع يساوى 10 على الوجهين فى القذفة الثانية }

نلاحظ أن الحوادث السابقة هى حوادث مستقلة . كما نلاحظ أن الاحتمال المطلوب هو :

$$P[(A_1 \cap B_2) \cup (B_1 \cap A_2)]$$

باستخدام النتيجة (١,١) نجد :

$$P[(A_1 \cap B_2) \cup (B_1 \cap A_2)] = P(A_1 \cap B_2) + P(B_1 \cap A_2)$$

وباستخدام التعريف (١,١٢) نجد أن :

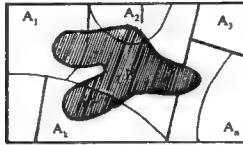
$$= P(A_1) P(B_2) + P(B_1) P(A_2)$$

$$= \left(\frac{4}{36}\right) \cdot \left(\frac{3}{36}\right) + \left(\frac{3}{36}\right) \cdot \left(\frac{4}{36}\right)$$

$$= 0,0185$$

### (١,٨) نظرية بيز أو ( قاعدة بيز ) Bayes rule

لنفرض أن الحوادث  $A_1, A_2, \dots, A_n$  تشكل تجربة منتهية لفضاء العينة  $S$  المتعلق بتجربة معينة ( هذا يعنى أن هذه الحوادث متنافية تبادليا . كما أن اتحادها يمثل الحادث الأكيد  $S$  ) ، ولنفرض أنه يتعلق بنفس التجربة حادث آخر  $B$  احتمال  $P(B) > 0$  كما هو موضح على الشكل (١,٧)



الشكل (١,٧)

عندئذ ستكون الحوادث  $A_i \cap B$  متنافية تبادليا ، وذلك مهما يكن  $i = 1, \dots, n$ .  
نلاحظ أن

$$B = S \cap B = [A_1 \cup A_2 \cup \dots \cup A_n] \cap B$$

$$= (A_1 \cap B) \cup (A_2 \cap B) \cup \dots \cup (A_n \cap B)$$

وباستخدام النظرية (١,١١) فإننا نجد أن :

$$P(B) = P(A_1) P(B/A_1) + P(A_2) P(B/A_2) + \dots + P(A_n) P(B/A_n)$$

من ناحية أخرى نعلم أن :

$$P(A_i/B) = \frac{P(A_i \cap B)}{P(B)}$$

وبالتعويض نجد أن :

$$P(A_i/B) = \frac{P(B/A_i) \cdot P(A_i)}{P(A_1) P(B/A_1) + \dots + P(A_n) P(B/A_n)} > 0 \quad 1 \leq i \leq n$$

## نظرية (١,١٣) Bayes rule

لتكن  $A_1, \dots, A_n$  مجموعة من الحوادث المشكلة لتجزئة فضاء العينة  $S$  المتعلق بتجربة معينة ، وحيث أن  $P(A_i) \neq 0$  من أجل جميع قيم  $i = 1, \dots, n$  لنفرض أن  $B$  حادثاً يتعلق بنفس التجربة ويحقق الشرط  $P(B) \neq 0$  ، عندئذ نجد أنه من أجل جميع قيم  $i = 1, \dots, n$

$$P(A_i/B) = \frac{P(B/A_i) \cdot P(A_i)}{\sum_{j=1}^n P(A_j) P(B/A_j)}$$

مثال (١,٢٩)

ثلاثة صناديق تحوى كرات حمراً ، وبيضاً ، وزرقاً كما هو موضح في الجدول التالى :

الصندوق الثالث	الصندوق الثاني	الصندوق الأول	اللون
3	4	2	عدد الكرات الحمراء
4	1	3	عدد الكرات البيضاء
3	3	5	عدد الكرات الزرقاء

سحبنا صندوقاً بصورة عشوائية من بين الصناديق الثلاثة ، ثم سحبنا منه كرة بصورة عشوائية أيضاً فكانت حمراء . ما هو احتمال أن يكون الصندوق المسحوب هو الأول ؟

الحل

نفرض الحوادث التالية :

$A_i$  = [ الصندوق المسحوب هو الصندوق ذو الرقم  $i$  وحيث إن  $i = 1, 2, 3$  ]

$B$  = [ الكرة المسحوبة من الصندوق حمراء ]

فيكون المطلوب حساب  $P(A_i/B)$

نلاحظ أن  $A_1, A_2, A_3$  تمثل ثلاث حوادث متنافية تبادلياً ، وأن اتحادها هو  $S$  ( لأنه لا بد من اختيار صندوق ، ولا يمكن أن نختار صندوقين دفعة واحدة ، كما أن

$$P(A_1) = P(A_2) = P(A_3) = \frac{1}{3}$$

كذلك فإن سحب كرة حمراء أى وقوع الحادث B يمكن أن يعم من الصندوق الأول ،  
أو الثانى ، أو الثالث أى إن B تقطع الحوادث  $A_1, A_2, A_3$  ، كما أن  $P(B) \neq 0$   
وبحسب النظرية (١,١٣) نجد أن :

$$P(A_1/B) = \frac{P(B/A_1) P(A_1)}{P(B/A_1) P(A_1) + P(B/A_2) P(A_2) + P(B/A_3) P(A_3)}$$

لذلك فإن :

$$P(B/A_1) = \frac{2}{10}, P(B/A_2) = \frac{4}{8}, P(B/A_3) = \frac{3}{10}$$

وبالتعويض فى العلاقة السابقة نجد أن

$$P(A_1/B) = \frac{\frac{2}{10} \cdot \frac{1}{3}}{\frac{2}{10} \cdot \frac{1}{3} + \frac{4}{8} \cdot \frac{1}{3} + \frac{3}{10} \cdot \frac{1}{3}} = \frac{2}{10} = 0.2$$



## تمارين محلولة

### تمرين (١)

كم طريقة يمكن أن يجلس بها ثمانية أشخاص على مقعد يتسع لثلاثة أشخاص فقط ؟

#### الحل

لنتصور الأمكنة الثلاثة a, b, c كما هو موضح □ □ □ نلاحظ أن المكان a يمكن شغله بثاني طرق من قبل كل شخص . وفي حالة ملء المكان a بأحد الأشخاص الثانية ، يبقى لدينا سبعة أشخاص وعندئذ يمكن اشغال المكان b بسبع طرق مختلفة ، وبعد شغل المكانين الأول والثاني بشخصين يبقى لدينا ستة أشخاص وست طرق لإشغال المكان c ، وعليه فإن عدد الطرق المطلوبة هو :

$$8.7.6 = 336$$

### تمرين (٢)

كم عددا مؤلفا من أربعة أرقام يمكن تكوينها من الأرقام العشرة التالية 0, 1, 2, ..., 9 ، وذلك في الحالتين التاليتين ؟

١ — التكرار ممكن

٢ — التكرار غير ممكن

#### الحل

الحالة الأولى : إذا كان تكرار أى رقم مسموح به في عملية التشكيل ، في هذه الحالة يمكن أن يكون الرقم الأول أى رقم من 0, 1, 2, ..., 9 أما الأرقام الثاني والثالث والرابع

فيمكن أن يكون كلاً منها أحد الأرقام العشرة المفروضة . فيكون مجموع الأعداد المطلوبة هو

$$9.10.10.10 = 9000$$

عددًا .

الحالة الثانية : أما إذا كان تكرار أى رقم غير ممكن ، فإننا نلاحظ أن العدد المطلوب من الأعداد هو  $9.9.8.7 = 4536$  عددًا .

تمرين (٣)

سحبنا ثلاث بطاقات من مجموع خمسين بطاقة . فإذا فرضنا أنه لا أهمية لترتيب عملية السحب ، فما هو عدد نقاط فضاء العينة في هذه التجربة ؟

الحل

نحسب النظرية (١,٤) فإن فضاء العينة S يحتوى على عدد من النقاط مساو لـ

$$\binom{50}{3} = \frac{50!}{(50-3)! \times 3!} = \frac{50!}{47! \times 3!} = \frac{(50)(49)(48)}{3 \times 2 \times 1} = 19600$$

تمرين (٤)

نود شراء آلة لصنع الحزير من إحدى خمس شركات . فبكم طريقة يمكن اختيار ثلاث من هذه الشركات ؟

الحل

حسب النظرية (١,٧) ، نجد أن عدد الطرق التي يمكن أن نختار بها ثلاث شركات من الخمس المفروضة هو :

$$\binom{5}{3} = \frac{5!}{3!2!} = \frac{5 \cdot 4}{2} = 10$$

تمرين (٥)

إذا علمت  $A \subset B$  . فبين أن  $P(A) \leq P(B)$

الحل

من الواضح أن :

$$P(B) = P(B \cap S) = P[B \cap (A \cup A')]$$

$$= P[(B \cap A) \cup (B \cap A')]$$

نلاحظ أن الحادتين  $B \cap A$  ،  $B \cap A'$  متنافيان ، لأنهما جزئان من الحادتين المتنافيين  $A$  . وبحسب النتيجة (١,١) نجد أن :

$$P(B) = P(B \cap A) + P(B \cap A')$$

وباعتبار أن

$$B \cap A = A \quad \text{إذن} \quad B \supset A$$

وهكذا فإن

$$P(B) = P(A) + P(B \cap A')$$

$$P(B) \geq P(A) \quad \text{إذن} \quad P(B \cap A') \geq 0$$

وبما أن

تمرين (٦)

بفرض أن احتمال سحب بطاقة رقمها مؤلف من ستة أعداد مجموع أول ثلاثة منها يساوى مجموع الثلاثة الأخر هو  $P = 0.05525$  . ابحث عن احتمال الحصول على بطاقة من هذا النوع ، وذلك من بين بطاقتين إذا كانت :

(١) كلا البطاقتين لهما أرقام متتالية .

(٢) إحدى البطاقتين مسحوبة بصورة عشوائية مستقلة عن البطاقة الثانية .

الحل

بفرض  $A$  = [ البطاقة المسحوبة الأولى لها مجموعات متساوية ]

$B$  = [ البطاقة المسحوبة الثانية لها مجموعات متساوية ]

أولا :

$$P(A \cup B) = P(A) + P(B) = 2 \cdot P(A) = 2 \cdot P = 0.1105$$

ثانيا : باستخدام النظرية (١,٩) والتعريف (١,١٢) نجد أن :

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = 2P - P^2 = 0.10744$$

## تمرين (٧)

يفرض أن الحادث  $A$  يمثل ظهور وجه الصورة ، وأن الحادث  $B$  يمثل ظهور أحد الوجهين 2 أو 4 وذلك لدى إلقاء حجر نرد وقطعة نقود دفعة واحدة . أوجد :

$$P(A \cap B) = 1$$

$$P(A/B) = 2$$

$$P(\bar{A} \cup \bar{B}) = 3$$

## الحل

أولاً : نلاحظ أن الحادث  $A \cap B$  يقع فيما لو ظهرت صورة على قطعة النقود ، ولم يظهر أحد الوجهين 2 أو 4 على حجر النرد لدى إلقائه وأن الحادثين  $A, B$  مستقلان .

ويمكن بحسب النتيجة (١,١) والنظرية (١,١٠) ، أن نكتب :

$$P(A \cap B) = P(A) \cdot P(B) = P(A) [1 - P(B)]$$

$$= \frac{1}{2} \left[ 1 - \left( \frac{1}{6} + \frac{1}{6} \right) \right]$$

ومنه :

$$P(A \cap B) = \frac{1}{3}$$

ثانياً : أما الحادث  $(A/B)$  فيمثل ظهور  $A$  بمعلومية  $B$  ، وبما أن  $B, A$  مستقلين

فإن :

$$P(A/B) = P(A) = \frac{1}{2}$$

ثالثاً : وأخيراً فإن الحادث  $A \cup B$  يمثل ظهور الكتابة على قطعة النقد ، أو أحد

الأعداد 1,3,5,6 أو كلاهما معا . ونعلم حسب النظرية (١,٩) أن :

$$P(A \cup B) = P(\bar{A}) + P(\bar{B}) - P(A \cap B)$$

ثم باستخدام النظرية (١,١٠) نجد أن :

$$P(\bar{A}) = 1 - P(A) = \frac{1}{2}, P(\bar{B}) = \frac{4}{6} = \frac{2}{3}$$

ومنه :

$$P(A \cup B) = \frac{1}{2} + \frac{2}{3} + \frac{1}{2} + \frac{2}{3} \\ = 0.8333$$

تمرين (٨)

إذا علمت أن الحادثين  $A, B$  متنافيان  $(A \cap B = \emptyset)$  ، وأن احتماليهما غير معدومين ، فهل يكون  $B, A$  مستقلين ؟

الحل

من التناقض نجد أنه وقع  $A$  فلا يمكن أن يقع  $B$  ، وهذا يعني أن  $P(B/A) = 0$  ، وبما أننا فرضنا أن  $0 < P(B)$  إذا  $P(B/A) \neq P(B)$  وهذا يعني أن :

$$P(A \cap B) \neq P(A) P(B/A) \neq P(A) P(B)$$

وحسب النظرية (١,١١) نجد أن الحادثين  $B, A$  غير مستقلين .

تمرين (٩)

عند سحب ورقتين من ورق اللعب ما هو احتمال الحصول على ملكة وتسعة ؟

الحل

لنرمز للحوادث  $C, B, A$  كالآتي :

$A = \{ \text{الورقتان المسحوبتان ملكة وتسعة} \}$

$B = \{ \text{الورقة الأولى ملكة والثانية تسعة} \}$

$C = \{ \text{الورقة الأولى تسعة والثانية ملكة} \}$

نلاحظ أن  $B \cup C = A$  ، وأن  $C, B$  حادثين متنافيين ، لحساب  $P(B)$  ،  $P(C)$  نعرف الحوادث الآتية :

$B_1 = \{ \text{الورقة المسحوبة الأولى ملكة} \}$

$B_2 = \{ \text{الورقة المسحوبة الثانية ملكة} \}$

$B_3$  = الورقة المسحوبة الأولى تسعة |

$B_4$  = الورقة المسحوبة الثانية تسعة |

واعتماداً على التعريف (١,١١) نجد أن :

$$P(B) = P(B_1 \cap B_2) = P(B_1) P(B_2/B_1) = \frac{4}{52} \cdot \frac{4}{51}$$

$$P(C) = P(B_3 \cap B_4) = P(B_3) P(B_4/B_3) = \frac{4}{52} \cdot \frac{4}{51}$$

وحسب النتيجة (١,١) نجد :

$$P(A) = P(B) + P(C) = \frac{8}{663} = 0.01206$$

تمرين (١٠)

كتبنا على بطاقات خاصة الأحرف المكونة لكلمة عندان ، والبالغ عددها خمس بطاقات ، ثم خلطناها خلطاً جيداً ، ثم سحبنا هذه البطاقات الواحدة تلو الأخرى . ما هو احتمال أن تشكل الأحرف المسحوبة كلمة عندان ؟

الحل

نلاحظ أن كلمة عندان تتألف من أربعة حروف هي العين والنون والذال والألف وقد ورد حرف العين مرة واحدة ، والنون مرتين والذال مرة ، والألف مرة أيضاً ، لذلك فإن :

$$P = \left(\frac{1}{5}\right) \cdot \left(\frac{1}{4}\right) \cdot \left(\frac{2}{3}\right) \cdot \left(\frac{1}{2}\right) \cdot \left(\frac{1}{1}\right) = \frac{1}{60} = 0.0167$$

تمرين (١١)

احسب  $P(A \cup B)$  ،  $P(A \cap B)$  إذا علمت أن  $P(A) = P(B) = \frac{1}{2}$  وأن

احسب

$$P(A/B) = \frac{2}{3}$$

الحل

نلاحظ من النظرية (١,١١) أن :

$$P(A \cap B) = P(B) \cdot P(A/B) = \frac{1}{2} + \frac{2}{3} = \frac{1}{3}$$

وحسب النظرية (١,٩) نجد أيضا أن :

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = \frac{1}{2} + \frac{1}{2} - \frac{1}{3} = \frac{2}{3}$$

تمرين (١٢)

راميان يصيب أحدهما الهدف باحتمال قدره 0.8 ويصيب الآخر نفس الهدف باحتمال قدره 0.7 . ماهو احتمال إصابة الهدف إذا أطلق منهما طلقة من بندقيته ؟

الحل

$$\begin{aligned} A &= \{ \text{أصاب الرامي الأول الهدف} \} \\ B &= \{ \text{أصاب الرامي الثاني الهدف} \} \end{aligned}$$

لنرمز بـ

باستخدام النظرية (١,٩) نجد أن :

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

ومن الفرض لدينا :

$$P(A) = 0.8, P(B) = 0.7$$

وحسب التعريف (١,١٢) نجد أن :

$$P(A \cap B) = P(A) \cdot P(B) = (0.8) \cdot (0.7)$$

وهكذا نجد أن :

$$P(A \cup B) = 0.8 + 0.7 - (0.8) \cdot (0.7) = 0.94$$

تمرين (١٣)

رجل وزوجته عمرهما على الترتيب 45,60 عاما . واحتمال وفاة كل منهما خلال الأعوام العشرة القادمة هو على الترتيب 0.45, 0.7 فإذا علمنا أن أحدهما قد توفي ، فما هو احتمال أن يكون الرجل ؟

## الحل

إذا رمزنا للحادثين  $B, A$  بالرمزين :

$A = \{ \text{وفاة الرجل خلال الفترة المحددة} \}$

$B = \{ \text{وفاة أحد الشخصين} \}$

فيكون المطلوب حساب  $P(A/B)$  . وحسب النظرية (١,١١) نجد أن :

$$P(A/B) = \frac{P(A \cap B)}{P(B)}$$

غير أن الحادث  $A \cap B$  يمثل وفاة أحد الزوجين ووفاة الرجل بنفس الوقت ، أى بوفاة الرجل وبقاء الزوجة على قيد الحياة . وهذا الاحتمال يساوى حسب النظرية (١,١٠)

$$(0.7)[1-0.45] = 0.385$$

ثم أن :

$$P(B) = (0.7) + (0.45) - 2(0.7)(0.45) = 0.52$$

$$P(A/B) = \frac{0.385}{0.52} = 0.74038 \quad \text{وبالعودة إلى العلاقة (*) نجد أن :}$$

## تمرين (١٤)

كتب عدنان ثلاث رسائل إلى أصدقائه في الولايات المتحدة ، ووضع كلا منها في ظرف . واختلطت الظروف الثلاثة قبل أن يعنونها ، بعد ذلك عنون هذه الرسائل بصورة عشوائية . ما هو احتمال أن يكون قد كتب على ظرف واحد على الأقل العنوان الصحيح ؟

## الحل

فترض أن  $A_k = \{ \text{كتب عدنان العنوان الصحيح على ظرف ذو الرقم } k \}$

حيث إن  $k = 1, 2, 3$  ، فيكون المطلوب عندئذ حساب  $P(A_1 \cup A_2 \cup A_3)$  باستخدام النظرية (١,٩) من أجل ثلاث حوادث نجد أن :

$$P(A_1 \cup A_2 \cup A_3) = [P(A_1) + P(A_2) + P(A_3)] - [P(A_1 \cap A_2) + P(A_1 \cap A_3) + P(A_2 \cap A_3)] + P(A_1 \cap A_2 \cap A_3)$$



غير أن :

$$P(A_k) = \frac{1}{3} \quad ; \quad k = 1, 2, 3$$

$$P(A_1 \cap A_2) = P(A_1) \cdot P(A_2/A_1) = \frac{1}{3} \cap \frac{1}{2} = \frac{1}{6}$$

$$P(A_1 \cap A_2) = \frac{1}{6}, P(A_2 \cap A_3) = \frac{1}{6}$$

وباستخدام النظرية (١,١٢) نجد أن :

$$P(A_1 \cap A_2 \cap A_3) = P(A_1) \cdot P(A_2/A_1) \cdot P(A_3/A_1 \cap A_2)$$

$$\frac{1}{3} \cdot \frac{1}{2} \cdot \frac{1}{1} = \frac{1}{6}$$

أخيرا وبالتعويض في العلاقة (\*) نستنتج الاحتمال المطلوب :

$$P(A_1 \cup A_2 \cup A_3) = \left[ \frac{1}{3} + \frac{1}{3} + \frac{1}{3} \right] - \left[ \frac{1}{6} + \frac{1}{6} + \frac{1}{6} \right] + \frac{1}{6}$$

$$= \frac{2}{3} = 0.666$$

### تمرين (١٥)

لعب أحمد وسعيد عشر مباريات في كرة الطاولة . كسب أحمد خلالها 4 مباريات ، كما كسب سعيد 3 مباريات وتعادلا في 3 مباريات واتفقا على اللعب ثلاث مباريات أخرى ، ما هو :

- ١) احتمال أن يربح أحمد جميع المباريات ؟
- ٢) احتمال أن يكسب أحمد وسعيد بالتبادل ؟
- ٣) احتمال أن تنتهي مباراتان بالتعادل ؟

### الحل

لنسمى الحوادث :

A = [ كسب أحمد في مباراة ]

B = [ كسب سعيد في مباراة ]

C = [ تعادل أحمد وسعيد في مباراة ]

من شروط اللعب الأول نجد أن :

$$P(A) = \frac{4}{10}, P(B) = \frac{3}{10}, P(C) = \frac{3}{10}$$

**أولاً :** احتمال أن يربح أحمد المباريات الثلاث يساوى احتمال أن يربح المباراة الأولى والثانية والثالثة .

$$P(AAA) = P(A) \cdot P(A) \cdot P(A) = \left(\frac{4}{10}\right)^3 = 0.064$$

**ثانياً :** ( إذا كسب أحمد ، ثم كسب سعيد ، ثم كسب أحمد ) أو ( كسب سعيد ، ثم كسب أحمد ، ثم كسب سعيد ) فهذا يعنى أنهما تعادلا بالتبادل وهذا الاحتمال يحسب بالشكل التالى :

$$P[A \cap B \cap A \cup (B \cap A \cap B)] = P(A \cap B \cap A) + P(B \cap A \cap B) =$$

$$P(A) \cdot P(B) \cdot P(A) + P(B) \cdot P(A) \cdot P(B) = 0.84$$

وبالتعويض نجد أن الاحتمال = 0.84

**ثالثاً :** لحساب هذا الاحتمال نلاحظ أن إنهاء مباراتين بالتعادل يعنى انتهاء المباراتين الأولى والثانية ، أو الأولى والثالثة ، أو الثانية والثالثة بالتعادل ، وهذا يعنى أن الاحتمال المطلوب هو :

$$P[(C \cap C \cap \bar{C}) \cup (C \cap \bar{C} \cap C) \cup (\bar{C} \cap C \cap C)] =$$

$$3P(C) \cdot P(C) \cdot P(\bar{C}) = 3 \cdot \frac{3}{10} \cdot \frac{3}{10} \cdot \frac{7}{10} = 0.189.$$

**تمرين (١٦)**

بفرض أن إنتاج ثلاثة أنواع من آلات حياكة الثياب هي على التوالي : 0.2, 0.3, 0.5 من الإنتاج الكلى لمصنع نسيج ، فإذا كانت نسبة الإنتاج المعاب لهذه الآلات هي على التوالي : 0.03 , 0.04 , 0.05 ، وإذا اخترنا قطعة بصورة عشوائية من الإنتاج العام ، ووجدنا أنها ذات عيب ، فما هو احتمال أن تكون هذه القطعة من حياكة الآلة ذات الإنتاج 0.5 ؟

## الحل

لنرمز للآلات الثلاث بالرموز C, B, A على الترتيب ، وبالرمز D للقطعة المسحوبة  
إن كانت معابة ، فيكون المطلوب حسب النظرية (١٣, ١) :

$$P(A/D) = \frac{P(D/A) \cdot P(A)}{P(D/A) \cdot P(A) + P(D/B) \cdot P(B) + P(D/C) \cdot P(C)}$$

وهكذا نجد أن :

$$P(A/D) = \frac{(0.03) (0.5)}{(0.03) (0.5) + (0.04) (0.3) + (0.05) (0.2)}$$

## تمرين (١٧)

مجموعتان من البضائع المصنعة ، عناصر المجموعة الأولى جميعا من النوع الجيد ،  
أما عناصر المجموعة الثانية فربعضها معاب . اخترنا مجموعة بصورة عشوائية من بين  
المجموعتين ، ثم سحبنا عنصراً منها فظهر أنه جيد ، ما هو احتمال أن نسحب عنصراً آخر  
من المجموعة المسحوبة فيظهر أنه ذو عيب ، إذا علمت أننا أعدنا العنصر المسحوب في  
المرّة الأولى إلى نفس المجموعة المسحوبة ؟

## الحل

لنسمى الحوادث التالية :

$A_1$  = [ المجموعة المختارة هي المجموعة الثانية ]

$A_2$  = [ المجموعة المختارة هي المجموعة الأولى ]

$B_1$  = [ العنصر المسحوب الأول من النوع الجيد ]

من شروط المسألة نجد أن :

$$P(A_1) = P(A_2) = \frac{1}{2}$$

$$P(B/A_1) = \frac{3}{4} , P(B/A_2) = 1$$

وحسب العلاقة :

$$P(B) = P(B/A_1) \cdot P(A_1) + P(B/A_2) \cdot P(A_2)$$

نجد أن :

$$P(B) = \frac{3}{4} \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} = 0.875$$

وبعد أول سحب فإن احتمال أن تحوى المجموعة المختارة عنصراً غير جيد ( معاب ) بحسب من النظرية (١,١٣) :

$$P(A_1/B) = \frac{P(A_1) \cdot P(B/A_1)}{P(A)}$$

$$P(A_1/B) = \frac{\frac{1}{2} \cdot \frac{3}{4}}{0.875} = \frac{3}{7} = 0.42857$$

أما احتمال أن تحتوى المجموعة المختارة على عنصر جيد فهو :

$$P(A_2/B) = \frac{4}{7}$$

لنرمز للحدث | العنصر الثانى المسحوب غير جيد ( معاب ) | D  
نلاحظ أن :

$$P(D/A_1) = \frac{1}{4}$$

$$P(D/A_2) = 0$$

ولذلك فإن الاحتمال المطلوب بحسب بالعلاقة :

$$P(D) = P(D/A_1) \cdot P(A_1) + P(D/A_2) \cdot P(A_2)$$

وهذا الاحتمال يساوى :

$$= \frac{3}{7} \cdot \frac{1}{4} = \frac{3}{28} = 0.107142857$$

تمرين (١٨)

يُنتج مصنع للأتابيب الإلكترونية نوعاً من اللمبات ، وقد لوحظ أن متوسط اللمبات المعيبة في إنتاج آلة معينة هو 20% ما هو احتمال :

- (١) وجود لمبتين معابتين عند اختيار عشرة بصورة عشوائية ؟
- (٢) وجود أكثر من لمبتين معابتين عند اختيار عشرة بصورة عشوائية ؟
- (٣) وجود أكثر من خمس لمبات معابة عند اختيار عشرة بصورة عشوائية ؟

الحل

لنرمز للحوادث :

A = | عدد اللمبات المعابة اثنتين |

B = | عدد اللمبات المعابة أكثر من لمبتين |

C = | عدد اللمبات المعابة أكثر من خمس |

نلاحظ أن :

$$1) P(A) = \binom{10}{2} (0.2)^2 \cdot (0.8)^8 = 0.0302$$

$$2) P(B) = 1 - (\text{أن يوجد اثنتين أو أقل معالجين})$$

$$= 1 - \left( \binom{10}{0} (0.2)^0 \cdot (0.8)^{10} + \binom{10}{1} (0.2)^1 \cdot (0.8)^9 \right. \\ \left. - \binom{10}{2} (0.2)^2 (0.8)^8 \right)$$

$$= 1 - 0.1074 - 0.2681 - 0.0302 = 0.594$$

وأخيرا فإن :

$$3) P(C) = P[\text{عدد اللمبات المعابة} \leq 6] + P[\text{عدد اللمبات المعابة} = 7] + \\ P[\text{عدد اللمبات المعابة} = 8] + P[\text{عدد اللمبات المعابة} = 9] + \\ P[\text{عدد اللمبات المعابة} = 10]$$

$$= \binom{10}{6} (0.2)^6 (0.8)^4 + \binom{10}{7} (0.2)^7 (0.8)^3 + \\ \binom{10}{8} (0.2)^8 (0.8)^2 + \binom{10}{9} (0.2)^9 (0.8)^1 + \\ \binom{10}{10} (0.2)^{10} (0.8)^0 \\ = 0.00637$$

### تمرين (١٩)

تتألف مجموعة عناصر من مائة عنصر تخضع للفحص المختبري الجزئي ، وشرط على عدم صلاحية هذه المجموعة إذا توافر فيها عنصر غير سليم على الأقل بين كل خمسة عناصر مراقبة . ما هو احتمال أن تكون مجموعة العناصر السابقة غير صالحة إذا احتوت على 5% من العناصر غير السليمة ؟

### الحل

بفرض | المجموعة صالحة |  $\bar{A}$  عندئذ يكون المطلوب حساب  $P(A)$  ومن المعلوم

$$P(\bar{A}) = 1 - P(A) = 1 - q \quad \text{أن :}$$

لنسحب  $q$  ، إذا رمزنا للحوادث ( العنصر المراقب ذى الرقم  $K$  سليم  $A_k$  ) .  
عندئذ نجد أن :

$$A = A_1 \cap A_2 \cap \dots \cap A_5$$

والملاحظ أن :

$$P(A_1) = \frac{95}{100}$$

لأن المائة عنصر تحوى 5 عناصر غير سليمة . وبعد تحقق الحادث  $A_1$  يبقى 99 عنصرا من  
بينها أربع وتسعون عنصرا سليما ولذلك فإن :

$$P(A_2/A_1) = \frac{94}{99}$$

وبنفس الطريقة نجد أن :  $P(A_2/A_1 \cap A_2 \cap A_3) = \frac{92}{97}$  ،  $P(A_3/A_1 \cap A_2) = \frac{93}{98}$  .

$$P(A_3/A_1 \cap A_2 \cap A_3 \cap A_4) = \frac{91}{96}$$

وباستخدام النظرية (١,١٢) فإننا نجد أن :

$$q = P(A) = P(A_1 \cap A_2 \cap \dots \cap A_5) = P(A_1) \cdot P(A_2/A_1) \cdot$$

$$P(A_3/A_1 \cap A_2) \cdot P(A_4/A_1 \cap A_2 \cap A_3) \cdot P(A_5/A_1 \cap A_2 \cap A_3 \cap A_4)$$

$$q = 0.77$$

$$P(A) = 0.23$$

## تمارين عامة

(١) عين عناصر فضاءات العينة التالية :

أ — مجموعة الأعداد الصحيحة المحصورة بين الواحد والخمسين والتي تقبل القسمة على سبعة .

ب —  $S = \{x | x^2 + x - 6 = 0\}$

ج — مجموعة النتائج الممكنة عند إلقاء حجر نرد وقطعة نقود في آن واحد

د —  $S = \{x | 2x - 4 = 0, x > 5\}$

(٢) عين الحوادث المتساوية فيما يلي :

$A = \{1, 3\}$

$B = \{x | \text{يمثل الوجه الذى ظهر عند إلقاء حجر نرد } x\}$

$C = \{x | x^2 - 4x + 3 = 0\}$

$D = \{x | \text{هو عدد مرات الصورة التى ظهرت عند إلقاء ست قطع نقود } x\}$

(٣) ألقينا حجرى نرد متماثلين ومتوازيين أحدهما أخضر والآخر أحمر ، والمطلوب :

أ — تحديد عناصر فضاء العينة  $S$  .

ب — تحديد عناصر الحادث  $A$  بمجموع الوجهين اللذين ظهرا أقل من خمسة  $A = \{$

ج — تحديد عناصر الحادث  $B$   $\}$  ظهر على أحد الوجهين العدد ستة  $B = \{$

د — تحديد عناصر الحادث  $C$   $\}$  ظهر على الحجر الأخضر الوجه اثنان  $C = \{$

هـ — ارسم مخطط فين لتوضيح العلاقة بين الحوادث  $S, A, B, C$

(٤) تتألف تجربة من إلقاء قطعة نقود ثم إعادة هذا الإلقاء إذا ظهر وجه صورة . فإذا علمت أنه قد ظهر كتابة على القطعة الأولى الملقاة ، وأنا ألقينا بعد ذلك حجر نرد فالمطلوب :

أ — تحديد عناصر فضاء العينة  $S$  .

ب — تحديد نقاط العينة الموافقة للحدث  $A = \{$  ظهور عدد أقل من 4 على حجر

$A = \{$  النرد

ج — تحديد نقاط العينة الموافقة للحدث  $B = \{$  ظهور الكتابة مرتين  $B = \{$

(٥) اختير أربعة طلاب من جامعة الملك عبد العزيز بجدة من كليتى الهندسة والطب ،

ثم صنفوا بحسب الفرع الذى يتبعون إليه . أوجد عناصر فضاء العينة  $S$  باستخدام الرمز  $E$  لطلاب الهندسة ، والرمز  $M$  لطلاب الطب ، بفرض  $S_2$  يمثل عدد طلاب كلية الهندسة المختارين ، فالمطلوب تحديد عناصر فضاء العينة  $S_2$  .

(٦) بفرض أن فضاء العينة :

$S = \{ \text{صوديوم ، بوتاسيوم ، نيتروجين ، يورانيوم ، أكسجين ، زنك ، فضة} \}$   
وأن الحوادث :

$A = \{ \text{فضة ، صوديوم ، زنك} \}$

$B = \{ \text{صوديوم ، نيتروجين ، بوتاسيوم} \}$

$C = \{ \text{أكسجين ، يورانيوم ، زنك} \}$

فالمطلوب تحديد عناصر الحوادث التالية :

أ :  $\bar{A}$

ب :  $A \cup C$

ج :  $(A \cap B) \cup \bar{C}$

د :  $B \cap \bar{C}$

هـ :  $A \cap B \cap C$

(٧) إذا علمت أن الحادثين :

$$A = \{x | 1 < x < 9\}, B = \{y | y < 5\}$$

فأوجد الحادثين :

$$A \cup B, A \cap B$$

(٨) بكم طريقة يمكن أن يصطف خمسة أشخاص لدى صعودهم إلى الباص ؟

(٩) ما هو عدد التباديل المختلفة التى يمكن تشكيلها من كلمة « الإحصاء » ؟

(١٠) بكم طريقة يمكن أن يجلس أربعة أطباء وثلاثة مهندسين فى صف إذا كان عليهم أن يجلسوا بالتناوب ؟

(١١) كم لجنة مؤلفة من ثلاثة أشخاص يمكن تشكيلها من ضمن خمسة مهندسين وثلاثة أطباء فى الحالات التالية :

أ — بدون تحديد نوع اللجان .



- ب — إذا ضُمَّت كل لجنة مهندسين وطبياً .  
 ج — إذا ضُمَّت كل لجنة طبيباً ومهندسا ، وكان هناك طبيب معين موجود في كل لجنة .

(١٢) يخترى صندوق على ثلاث تفاحات حمراء وأربع خضراء وخمس صفراء . بكم طريقة يمكن سحب ست تفاحات من الصندوق إذا ضُمَّت تفاحتين من كل لون ؟

(١٣) سحبنا ورقتين من ورق اللعب المؤلف من 52 ورقة وذلك بصورة عشوائية ، ما هو احتمال أن تكون كلتا الورقتين أكبر من 2 وأقل من 9 ؟

(١٤) إذا علمت أن احتمال الحادثين A, B المستقلين هما على الترتيب

$$P(A) = 0.4, \quad P(B) = 0.5 \quad \text{فما هو :}$$

$$P(A \cup B) \text{ و } P(A)P(B)$$

(١٥) يوجد سيارتان مستقلتان لإطفاء الخرائق في مركز إطفاء البغدادية في مدينة حدة ، فإذا علمت أن احتمال أن تكون سيارة معينة منهما جاهزة لإطفاء حريق هو 0.99 فامضلوب :

أ — حساب احتمال ألا تكون أية سيارة منهما جاهزة لإطفاء الحريق عندما نحتاجها .

ب — حساب احتمال أن تكون سيارة حريق جاهزة عند احتياجنا لها .

(١٦) إذا علمت أن احتمال أن يبقى أحمد على قيد الحياة خلال العشرين عاما المقبلة هو 0.6 ، وأن احتمال أن تبقى زوجته على قيد الحياة خلال العشرين عاما المقبلة هو 0.9 ، فما هو احتمال ألا يبقى أحمد أو زوجته على قيد الحياة خلال نفس الفترة ؟

(١٧) إذا علمت أن احتمال أن يفتح عمار التلفزيون على برنامج معين للأطفال هو 0.4 ، واحتمال أن تفتح شقيقته التلفزيون على نفس البرنامج هو 0.5 ، وأن احتمال أن يفتح عمار التلفزيون على البرنامج إذا علم أن شقيقته قد فتحت على نفس البرنامج هو 0.7 فامضلوب حساب :

أ — احتمال أن يفتح كلاهما على نفس البرنامج .

ب — احتمال أن تفتح البنت على البرنامج علما أن عمار قد فتح عليه .

ج — احتمال أن يفتح أحدهما على الأقل على نفس البرنامج .

(١٨) يخوى كيس أربع كرات بيض وثلاث سود ، ويخوى كيس آخر ثلاث كرات بيض وخمس سود . سحبنا بصورة عشوائية كرة من الكيس الأول ووضعناها في الثاني . ما هو احتمال سحب كرة سوداء من الكيس الثاني ؟

(١٩) ثلاثة صناديق تحتوي على كرات ملونة موزعة على النحو التالي :

اللون	الصندوق الأول	الثاني	الثالث
أحمر	12	4	3
أبيض	9	11	4
أزرق	5	6	8

سحبنا صندوقاً بصورة عشوائية ، ثم سحبنا منه كرة بصورة عشوائية أيضاً ، فلاحظنا أنها حمراء . ما هو احتمال أن يكون الصندوق المسحوب هو الثاني ؟

## الفصل الثاني

### لمتغيرات العشوائية

- المفهوم العشوائي ■ التوزيع الاحتمالي المنقطع ■ توزيع الاحتمال المستمر
- التوزيعات التجريبية ■ توزيع الاحتمال المشترك ■ التوقع الرياضي ■ قوانين
- التوقع الرياضي ■ التوقعات الرياضية الخاصة (البابن-التغير) ■ خواص التباين
- نظرية تشيشف ■ نماذج محلولة ■ نماذج عامة .



## (٢,١) المتغير العشوائى

### The random variable

من الواضح أن التجربة الإحصائية هى عملية تؤدي إلى قياس أو ملاحظات ، وأن معظم التجارب ذات الأهمية التجريبية تقودنا إلى قياسات عديدة تتغير من نقطة عينة إلى أخرى ، وبالتالي فإن هذا القياس يدعى بالمتغير العشوائى .

### تعريف (٢,١) المتغير العشوائى :

المتغير العشوائى هو دالة عددية تأخذ قيما حقيقية تتحدد بوساطة كل عنصر من عناصر فضاء العينة  $S$  .

سنرمز للمتغير العشوائى بأحد الأحرف الكبيرة ..  $X, Y, Z$  ، أما القيم التى يأخذها هذا المتغير فى نقطة عينة معينة فنسرمز لها بأحد الحروف الصغيرة ...  $x, y, z$  . بفرض أن التجربة الإحصائية تمثل إحصاء عدد الصور التى نحصل عليها لدى إلقاء قطعة نقود مرتين . فمن الواضح عندئذ أن عدد الصور هذا ما هو إلا متغير عشوائى نرمز له بالرمز  $X$  ، وهذا المتغير يأخذ القيم 0,1,2 عند نقاط العينة الأربعة فى فضاء العينة :

$$S = \{ HH, HT, TH, TT \}$$

### مثال (٢,١)

يحتوى وعاء على ثلاث كرات حمراء وخمس بيض . ولنسحب كرتين من هذا الوعاء بصورة عشوائية على التتالى ، وبدون إعادة الكرة المسحوبة إلى الوعاء . من الواضح أنه إذا رمزنا بـ  $X$  لعدد الكرات الحمراء المسحوبة فى هذه التجربة وبالرمز  $x$  لقيم هذا المتغير العشوائى فإننا نجد أن :

الجدول (٢،١)

X	الحوادث البسيطة
3	RRR
2	RWR
2	WRR
2	RRW
1	RWW
1	WRW
1	WWR
0	WWW

## مثال (٢،٢)

في تجربة إلقاء حجرى نرد دفعة واحدة ، إذا رمزنا بـ  $Y$  لأكبر العددين اللذين ظهرا على وجهي هذين الحجرين ، أى أنه من أجل أى نقطة  $(x,y) \in S$  فإن  $Y(x,y) = \max(x,y)$  . عندئذ تكون قيم المتغير  $Y$

$$Y: 1,2,3,4,5,6$$

## مثال (٢،٣)

لنرمز لعدد القذفات اللازمة للحصول على وجه الصورة لأول مرة عند إلقاء قطعة نقود بالرمز  $Y$  من الواضح أن قيم  $Y$  هي .

$$Y: 1,2,3,4,5,\dots$$

إذا أمعنا النظر في المثالين (٢،١) ، (٢،٢) فإننا نجد أن فضاء العينة في كل منهما يتألف من عدد منته من النقاط ، أما فضاء العينة في المثال (٢،٣) فيتألف من عدد غير منته من نقاط العينة إلا أنه معدود .

## تعريف (٢،٢) متغير عشوائى منقطع Discrete random variable

إذا حوى فضاء عينة ما على عدد منته أو غير منته ولكنه معدود من نقاط العينة ، قلنا إن هذا الفضاء منقطع ، ونسمى المتغير العشوائى المعرف على هذا الفضاء بمتغير عشوائى منقطع .

نسأل الآن أنفسنا عن فضاء العينة الموافق لتجربة قياس المعدل اليومي لهطول المطر في مدينة جدة خلال شهر معين ، إذا استخدمنا ( مسطرة مدرجة ) لهذا القياس ، فإن معدل الهطول سيوافق نقطة على هذه المسطرة والتي يمكن النظر إليها على أنها نقطة من محور موجه . وفضاء العينة يمكن إذاً أن يكون أى نقطة من مجال على هذا المحور . ومن المعلوم أنه يوجد من النقاط في أى مجال من محور موجه مهما كان صغيراً عدداً غير معدود من النقاط .

### تعريف (٢,٣) متغير عشوائى مستمر Continuous random variable

إذا حوى فضاء العينة في تجربة معينة على عدد غير معدود وغير محدود من النقاط قلنا عنه أنه فراغ عينة مستمر . كما نسمى المتغير العشوائى المعروف على هذا النوع من الفراغات بمتغير عشوائى مستمر .

نلاحظ أن الأوزان والأطوال والقوى كلها أمثلة على المتغيرات العشوائية المستمرة ، وقياسات مثل هذه المتغيرات تشكل بدون انقطاع نقاط على محور موجه أو مجالات بين قياس وآخر .

### (٢,٢) التوزيع الاحتمالى المنقطع

#### Discrete Probability Distribution

إن العدد الممثل للوجه الذى يظهر عند إلقاء حجر نرد والذى نرمز له بالرمز  $X$  هو متغير عشوائى يأخذ القيم 1,2,3,4,5,6 باحتمالات متساوية . كل منها يساوى العدد  $\frac{1}{6}$  . كما نلاحظ أن المتغير العشوائى الممثل لعدد الصور التى نحصل عليها لدى إلقاء قطعة نقود أربع مرات متتالية سيأخذ القيمة 3 باحتمال قدرة  $\frac{1}{4}$  . ونلاحظ أيضاً .

في المثال (٢,٢) أن المتغير العشوائى  $Y$  يفترض القيم 1,2,3,4,5,6 باحتمالات موافقة قدرها

$$\frac{1}{36}, \frac{3}{36}, \frac{5}{36}, \frac{7}{36}, \frac{9}{36}, \frac{11}{36} \text{ على الترتيب :}$$

ومن المناسب في كثير من الحالات أن نمثل الاحتمال الموافق لقيمة معينة من قيم المتغير العشوائى المفروض على شكل دالة بهذه القيمة ، نرمز لها بالرمز  $f(x)$  وعلى هذا

نكتب  $f(x) = P[X = x]$  . فمثلا نجد في المثال (٢,٢) أن :

$$f(1) = P[X = 1] = P[(1,1)] = \frac{1}{36}$$

$$f(2) = P[X = 2] = P[(2,1), (2,2), (1,2)] = \frac{3}{36}$$

وهكذا ....

### تعريف (٢,٤) التوزيع الاحتمالي Probability distribution

تسمى الدالة  $f(x)$  بتوزيع الاحتمال للمتغير العشوائى المنقطع  $X$  إذا حققت الشروط الثلاثة التالية :

$$1) f(x) \geq 0$$

$$2) \sum_x f(x) = 1$$

$$P(x) = P[X = x]$$

وذلك لجميع قيم  $x$  .

### مثال (٢,٤)

أوجد توزيع الاحتمال للمتغير العشوائى  $Y$  الوارد في المثال (٢,٢) ؟

### الحل

من الواضح ( انظر المثال ١,٩ ) أن عدد نقاط العينة عند إلقاء حجرى النرد هو 36 نقطة ، وكما نعلم فإن قيم المتغير  $Y$  هى  $\{1,2,3,4,5,6\}$  ولذلك :

$$f(1) = P(Y = 1) = P[(1,1)] = \frac{1}{36}$$

$$f(2) = P(Y = 2) = P[\{(1,2), (2,2), (2,1)\}] = \frac{3}{36}$$

$$f(3) = P(Y = 3) = P[\{(1,3), (3,1), (3,3), (2,3), (3,2)\}] = \frac{5}{36}$$

$$f(4) = P(Y = 4) = P[\{(1,4), (4,1), (2,4), (4,2), (3,4), (4,3), (4,4)\}] = \frac{7}{36}$$



وأخيراً فإن :

$$f(5) = P(Y = 5) = \frac{9}{36}, f(6) = P(Y = 6) = \frac{11}{36}$$

يمكن تنظيم قيم المتغير العشوائى والاحتمالات الموافقة بجدول كالتالى :

Y	1	2	3	4	5	6	
f(y)	$\frac{1}{36}$	$\frac{3}{36}$	$\frac{5}{36}$	$\frac{7}{36}$	$\frac{9}{36}$	$\frac{11}{36}$	$\sum f(y_i) = 1$ $Y_i$

مثال (٢,٥)

لنبحث عن توزيع الاحتمال لعدد مرات الصورة التى نحصل عليها لدى إلقاء قطعة نقود متماثلة ومتوازنة ثلاث مرات .

الحل

بالرجوع إلى المثال (١,١٨) نجد أن عدد نقاط فضاء العينة S هو 8 وأن هذه النقاط متساوية في إمكانية وقوعها ، أما عدد أوجه الصورة التى نحصل عليها فيساوى :  
X: 0,1,2,3 ، كما أن الاحتمالات الموافقة لهذه القيم فهى :

$$f(0) = P[X = 0] = \left(\frac{0}{8}\right) = \frac{1}{8}$$

$$f(1) = P[X = 1] = \left(\frac{1}{8}\right) = \frac{3}{8}$$

$$f(2) = P[X = 2] = \left(\frac{2}{8}\right) = \frac{3}{8}$$

$$f(3) = P[X = 3] = \left(\frac{3}{8}\right) = \frac{1}{8}$$

والملاحظ فى هذا المثال أننا استخدمنا عدد الطرق التى يمكن بها تجربة n شيئا إلى مجموعتين : تحتوى إحداهما على x شيئا ، والثانية على n-x شيئا آخر ، وهذا العدد ما هو إلا التوفيقه  $\left(\frac{n}{x}\right)$  . وهكذا نجد أن جدول توزيع المتغير العشوائى X :

X	0	1	2	3	
f(x)	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$	$\sum f(x_i) = 1$ $X_i$

تعريف (٢,٥) التوزيع التراكمي Cumulative distribution

نعرف التوزيع التراكمي  $F(x)$  للمتغير عشوائي منقطع  $X$  له توزيع احتمال  $f(x)$  بالعلاقة :

$$F(x) = P[X \leq x] = \sum_{y \leq x} f(y)$$

مثال (٢,٦)

لنبحث عن التوزيع التراكمي للمتغير العشوائي  $X$  في المثال (٢,٥)

الحل

نلاحظ أن :

$$F(0) = f(0) = \frac{1}{8}$$

$$F(1) = f(0) + f(1) = \frac{1}{8} + \frac{3}{8} = \frac{1}{2}$$

$$F(2) = f(0) + f(1) + f(2) = \frac{7}{8}$$

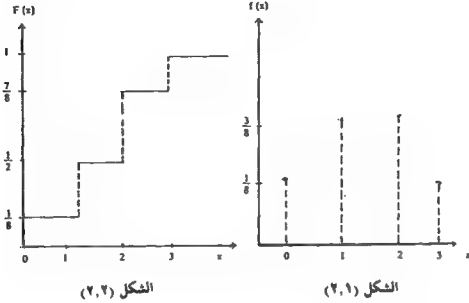
$$F(3) = f(0) + f(1) + f(2) + f(3) = 1$$

وهكذا نجد أن :

$$F(x) = \begin{cases} 0 & \text{for } x < 0 \\ \frac{1}{8} & \text{for } 0 \leq x < 1 \\ \frac{1}{2} & \text{for } 1 \leq x < 2 \\ \frac{7}{8} & \text{for } 2 \leq x \\ 1 & \text{for } 3 \leq x \end{cases}$$

تمثل بيانياً كلا من دالة توزيع الاحتمال ، ودالة التوزيع التراكمي .

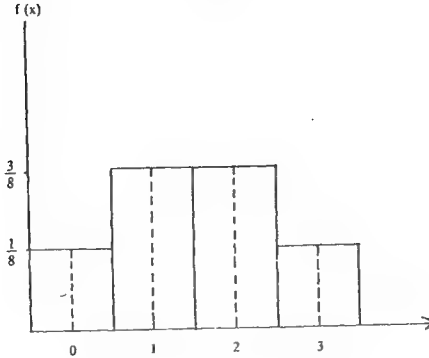
يبين الشكلين (١،٢) ، (٢،٢) على الترتيب الشكل البياني لدالة توزيع الاحتمال ، ولدالة التوزيع التراكمي .



### ملاحظة هامة

في كثير من الأحيان نقوم بدلاً من رسم النقاط  $(x, f(x))$  بإنشاء مستطيلات كما في الشكل (٢،٣) الموضح أدناه . هذه المستطيلات لها قواعد متساوية ومتمركزة في النقاط الممثلة لقيم المتغير العشوائي ، أما ارتفاعاتها فتساوي قيم  $f(x)$  الموافقة ويسمى الشكل (٢،٣) بالمضلع التكراري لـ  $X$  ويوضح لنا المضلع التكراري لـ  $X$  الوارد في المثال (٢،٥) والموضح على الشكل (٢،٣) أن  $P[X = x]$  ما هو إلا مساحة هذا المستطيل المتمركز في النقطة  $x$  ذلك لأن عرض المستطيل هو الواحد حتى ولو لم تكن المستطيلات ذات قواعد عرضها الواحد فإننا نقوم بتعديل ارتفاع هذه المستطيلات لتبقى مساحات المستطيلات في المضلع التكراري مساوية للاحتتمالات الموافقة لقيم المتغير  $X$  .

هذه الفكرة حول استخدام المساحات لتمثيل الاحتمالات ضرورية وهامة في دراسة دوال التوزيع للمتغيرات العشوائية المستمرة .



الشكل (٢,٣)

المضلع التكرارى

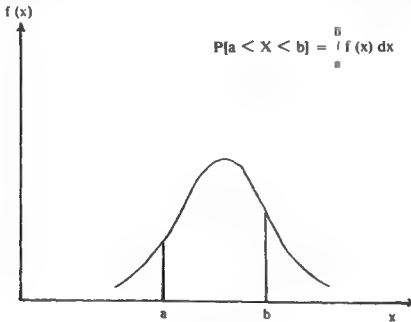
### (٢,٣) توزيع الاحتمال المستمر

#### Continuous propability distribution

كما أشرنا فى التعريف (٢,٣) إلى أن قياسات المتغيرات المستمرة تشكل بدون انقطاع على محور موجه أو مجالات بين قياس وآخر ، وبما أنه لا يمكننا تخصيص احتمال لكل نقطة عينة ، فلا بد من التفكير فى نموذج احتمالى مختلف عما رأيناه فى حالة المتغيرات المنقطعة . فلو عدنا إلى المثال (٢,٥) ، والشكل (٢,٣) الممثل للمضلع التكرارى فى هذا المثال ، وقمنا بزيادة عدد مرات إلقاء قطعة النقود بشكل كبير ، فعندئذ سيضيق عرض المجالات الجزئية ، وستتغير المظهر العام للمضلع التكرارى فى اتجاه التخلص من مظاهر عدم الانتظام . وعندما يصبح عدد القياسات كبيراً جداً ، وعرض المجالات الجزئية صغيراً جداً فسيظهر التكرار النسبى وكأنه عملياً منحنى لنعدل المساحة الكلية تحت المضلع التكرارى هذا لتصبح مساوية الواحد .

لنفرض الآن التحول العشوائى المستمر  $X$  يأخذ قيمة فوق محور الأعداد الحقيقية ضمن المجال  $(a, b)$  ، ولنوزع احتمالاً قدره الواحد على طول هذا المجال إلى حد كبير ، كما

يوزع شخص حفنة من الرمال ، بحيث توافق كل ذرة رمل قياساً من قياسات المجتمع . وستجتمع ذرات الرمل أو القياسات ، ونقول عندئذ أن كثافة الاحتمال في مثل هذه الأماكن أكبر منها في أماكن أخرى ، وسيكون لهذه الكثافة قيمة غير سالبة ( موجبة أو صفر ) في كل نقطة من نقاط المجال . فلو تصورنا أن هذه الكثافة تتغير من نقطة إلى أخرى وفق علاقة  $f(x)$  فإن منحنى الدالة  $f(x)$  هو منحنى الكثافة (density function) . ويظهر الشكل (٢،٤) توضيحاً لدالة الكثافة ، وبما أن المساحة أسفل المنحنى  $f(x)$  يجب أن تساوى الواحد ، لذا فإن المساحة أسفل المنحنى وفوق المجال  $(a,b)$  تساوى الواحد . وبشكل عام إذا كان  $X$  متغيراً عشوائياً مستمراً ، دالة كثافته الاحتمالية  $f(x)$  ، وإذا كان  $(a,b)$  مجالاً ما على المحور الحقيقي ، عندئذ يكون احتمال أن يأخذ  $x$  قيمة ضمن هذا المجال  $P[a < X < b]$  مساوياً إلى المساحة تحت منحنى الكثافة  $f(x)$  بين الحطين الشاقوليين  $x = a$  ،  $x = b$  أى إن :



$$P[a < X < b] = \int_a^b f(x) dx$$

الشكل (٢،٤)

### تعريف (٢،٦) دالة الكثافة الاحتمالية Probability density function

ليكن  $X$  متغيراً عشوائياً مستمراً معرفاً فوق المحور الحقيقي  $R$  . نسمى الدالة  $f(x)$  بدالة الكثافة الاحتمالية لهذا المتغير إذا حققت الشروط الثلاثة التالية :

$$1 - f(x) \geq 0 \quad \text{for all } x \in \mathbb{R}$$

$$2 - \int_{-\infty}^{+\infty} f(x) dx = 1$$

$$3 - P[a < X < b] = \int_a^b f(x) dx$$

## مثال (٧،٧)

ليكن  $X$  متغيراً عشوائياً مستمراً معرفاً على المجال  $(1,3)$  بالكثافة الاحتمالية :

$$f(x) = \frac{1}{2} \quad : \quad 1 < x < 3$$

$$= 0 \quad : \quad \text{فيما عدا ذلك}$$

لنرى إن كان  $f(x)$  يمثل دالة كثافة احتمالية أم لا ، ثم لنبحث عن  $P(2 < x < 2.5)$  . نلاحظ

$$f(x) = \frac{1}{2} > 0 \quad \text{فإن } x \text{ فوق المجال المذكور}$$

ثم إن

$$\int_{-\infty}^{+\infty} f(x) dx = \int_1^3 \frac{1}{2} dx = 1$$

وأخيراً فإن :

$$P(2 < X < 2.5) = \int_2^{2.5} \frac{1}{2} dx = \frac{1}{4}$$

## مثال (٧،٨)

بفرض أن الدالة :

$$f(x) = \frac{2(1+x)}{27} \quad : \quad 2 < x < 5$$

$$= 0 \quad : \quad \text{فيما عدا ذلك}$$

تمثل دالة كثافة لتغير عشوائى مستمر . احسب  $P(3 < X < 4)$  .

نلاحظ أن :

$$P[3 < x < 4] = \int_3^4 f(x)dx = \frac{2}{27} \int_3^4 (1+x) dx = \frac{1}{3}$$

### تعريف (٢,٧) التوزيع التراكمي The cumulative distribution

نقول بأن الدالة  $F(x)$  تمثل دالة التوزيع التراكمي لمتغير عشوائي مستمر  $X$  بالكثافة  $f(x)$  إذا حققت الدالة  $F(x)$  العلاقة الآتية :

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(u) du$$

من التعريف السابق نستنتج النتائج الهامة التالية :

$$1. \quad P(a < X < b) = F(b) - F(a)$$

$$2. \quad \frac{dF(x)}{dx} = f(x)$$

### مثال (٢,٩)

لنفتش في المثال (٢,٧) عن دالة التوزيع التراكمي  $F$  ، ثم لنحسب  $P[2 < X < 3]$

### الحل

من التعريف (٢,٧) نلاحظ أن :

(١) إذا كان  $X \leq 1$  فإن  $F(x) = 0$  لأنه لا توجد كثافة على المجال  $(-\infty, 1)$

(٢) أما إذا كان  $1 < X \leq 3$  فإن :

$$\begin{aligned} F(x) &= \int_{-\infty}^x f(u) du = \int_{-\infty}^1 f(u) du + \int_1^x f(u) du \\ &= \int_{-\infty}^1 0, du + \int_1^x \frac{1}{2} du = \frac{x-1}{2} \end{aligned}$$

(٣) وإذا كان  $x > 3$  فإننا نجد أن :

$$\begin{aligned}
 F(x) &= \int_{-\infty}^1 f(u)du + \int_1^3 f(u)du + \int_3^x f(u)du \\
 &= \int_{-\infty}^1 0 \cdot du + \int_1^3 \frac{1}{2} du + \int_3^x 0 \cdot du = 1
 \end{aligned}$$

وأخيرا فإننا نلاحظ أن :

$$F(x) = \begin{cases} 0 & \text{if } x \leq 1 \\ \frac{x-1}{2} & \text{if } 1 < x < 3 \\ 1 & \text{if } x \geq 3 \end{cases}$$

كذلك نلاحظ أن :

$$P(2 < x \leq 3) = F(3) - F(2) = \frac{3-1}{2} - \frac{2-1}{2} = 1 - \frac{1}{2} = \frac{1}{2}$$

### (٢,٤) التوزيعات التجريبية Empirical distribution

يمكن وصف مجموعة من القياسات بإحدى طريقتين : تدعى الأولى بالطريقة البيانية ، أما الثانية فتسمى بالطريقة العددية وهي بدئية ، حيث تقوم بوصف القياسات دون تحليلها . سنشرح فيما يلي الطريقة البيانية بشيء من التفصيل :

#### الطريقة البيانية :

يمثل الجدول (٢,٢) الموضوع أدناه معدلات ثلاثين طالبا من طلاب السنة الجامعية الأولى في إحدى جامعات المملكة . بالقاء نظرة سريعة على هذا الجدول يتبين لنا أن أقل معدل هو 46.7 ، وأعلى معدل هو 96.7 والسؤال المطروح هل : القياسات الثماني والعشرين الباقية تتوزع بصورة عادلة في المجال (46.7,96.7) ، أم أنها تقع بالقرب من أطراف المجال السابق ؟ ولكي نجيب على هذا السؤال نقوم بتقسيم المجال السابق إلى عدد من المجالات الجزئية المتساوية يتوقف عددها على عدد القياسات الموجودة ( عادة نختار من 5 إلى 20 مجالا جزئيا يتناسب عددها مع عدد القياسات التي ندرسها ) . فمثلا : يمكن أن نستخدم المجالات الجزئية (59.15,52.90) (65.40,59.15) (71.65,65.40) ... إلخ .



## الجدول (٢,٢)

الجدول (٢,٢) يبين معدلات 30 طالبا في إحدى جامعات المملكة

50.0	86.7	46.7	66.7	46.7
60.0	70.0	86.7	66.7	53.3
80.0	83.3	73.3	66.7	63.3
73.0	66.7	63.3	83.3	96.7
70.0	76.7	66.7	73.3	80.0
73.3	76.7	56.7	73.3	53.3

والمتمتع لهذه المجالات يلاحظ أن أى قياس لا يقع فوق نقاط التقسيم . لنصنف الآن القياسات الموجودة في الجدول (٢,٢) كلا في المجال الجزئى الذى يحويه ، يبين الجدول (٢,٣) هذا التصنيف :

## الجدول (٢,٣)

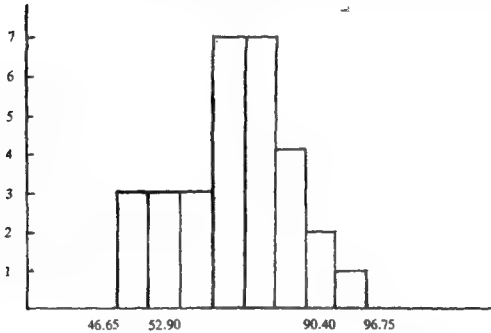
رقم المجال الجزئى	حدود كل مجال جزئى	عدد القياسات التي تقع في المجال الجزئى المقابل	عدد القياسات التي تقع في المجال الجزئى المقابل	نسبة القياسات الواقعة ضمن كل مجال جزئى إلى العدد الكلى للقياسات التى تصنفها
1	64.56 - 52.90	III	3	3/30
2	52.90 - 59.15	III	3	3/30
3	59.15 - 65.40	III	3	3/30
4	65.40 - 71.65	VII	7	7/30
5	71.65 - 77.90	VII	7	7/30
6	77.90 - 84.15	IV	4	4/30
7	48.15 - 90.40	II	2	2/30
8	90.40 - 96.75	I	1	1/30

المجموع

n = 30

1

فنسمى النسبة المذكورة في العمود الخامس بالتكرار النسبي (relative frequency) للمجال الجزئي الموافق . هذا ويمكن تمثيل الجدول الناتج بيانيا على شكل مضلع تكرارى حيث نقوم بإنشاء مستطيل ( فوق كل مجال جزئى ) ارتفاعه متناسب مع عدد القياسات الواقعة ضمن هذا المجال . يبين الشكل (٢,٥) مثل هذا التمثيل .



شكل (٢,٥) المضلع التكرارى

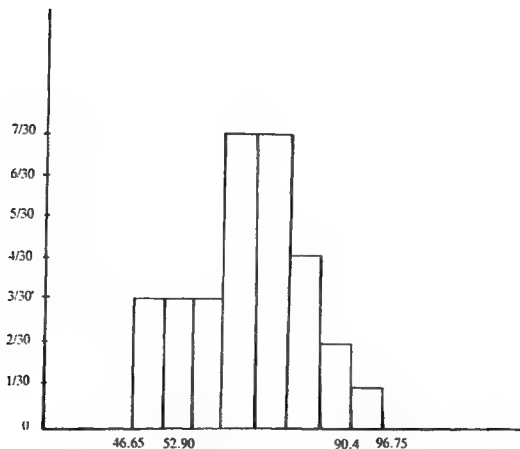
غالباً ما يكون من الأنسب تعديل المضلع التكرارى بحيث نختر وحدة التقسيمات على المحور الشاقولى . فإذا اعتبرنا أن ارتفاع كل مستطيل ننشئه فوق كل مجال جزئى مساوياً للتكرار النسبي الموافق لهذا المجال كما في الشكل (٢,٦) فإنه من النادر عندئذ التمييز بين المضلعين .

بالعودة إلى مضلع التكرار نلاحظ أن 14 طالباً قد حصلوا على معدلات أكبر من 71.65 . كما نلاحظ أيضاً أن هذه النسبة تساوى النسبة المثوية للمساحة الواقعة إلى يمين 71.65 . وبما أن 14 طالباً قد حصلوا كل منهم على معدلاً أكبر من 71.7 فإننا نقول أن هناك 14 فرصة من أصل 30 ، ونقول أيضاً بأن احتمال الحصول على مثل هذا المعدل هو

$$\frac{14}{30}$$

ولو عدنا إلى المجتمع ( مجموعة طلبة السنة الأولى ) الذى سجبنا منه هذا العدد من الطلاب فما هي نسبة الطلاب الذين نالوا معدلا أكبر من 71.7 ؟ نلاحظ أن الجواب لهذا السؤال يكمن فى حساب نسبة المساحة الكلية ( من مضلع التكرار النسبى الخاص بالمجتمع ) الواقعة إلى يمين 71.65 . ونظرا لعدم وجود مثل هذا المضلع أمامنا فإننا سنضطر للقيام باستنتاج . ولا بد قبل كل شيء من تقدير النسبة الحقيقية من المجتمع على أساس المعلومات التى تقدمها العينة ( درجات الطلاب الثلاثين ) . ومن المعقول أن يكون تقديرنا هنا هو  $\frac{14}{30}$  أو 47% .

لنفرض أننا نود التعبير عن احتمال أن يكون لطلاب انتقيناه من المجتمع بطريق الصدفة معدلا أكبر أو يساوى 71.7 .



الشكل (٢,٦)

يمكن أن نقول ( وبدون معرفة مضلع التكرار النسبي للمجتمع ) أن مضلع المجتمع مشابه لمضلع العينة ، وأن  $\frac{14}{30}$  من قياسات المجتمع على وجه التقريب قد تكون أكبر أو تساوى 71.7 . ومن الطبيعي أن نرتكب في عملية التقدير هذه خطأ . وسندرس مستقبلا كيف نضع حدوداً لمثل هذا الخطأ المرتكب في التقدير .

يدعى مضلع التكرار النسبي بالتوزيع التكرارى ، لأنه يبين الطريقة التى تتوزع فيها المعلومات على طول محاور الفواصل . والملاحظ أن المستطيلات المقامة فوق كل مجال جزئى بخاضعة لتفسيرين . فمن جهة تمثل نسبة الملاحظات الواقعة في مجال جزئى معين ، ثم إنه عند سحب قياس من القياسات وبشكل عشوائى يمكننا اعتبار مساحة المستطيل المقام فوق مجال جزئى معين ممثلة لاحتمال وقوع ذلك القياس ضمن هذا المجال الجزئى . والمهم في مضلع التكرار للعينة هو أن هذا المضلع يمدنا بمعلومات حول المضلع التكرارى للمجتمع ، والشئ المتوقع أن يكون مضلع التكرار للمجتمع والعينة متشابهين ، وأن تزداد درجة التشابه هذه كلما ازداد حجم العينة ، وعندما تزداد العينة لتصبح متطابقة مع المجتمع المدروس فعندئذ يتطابق المضلعان التكراريان .

### (٢،٥) توزيع الاحتمال المشترك Joint probability distribution

لقد اقتصرت دراستنا للمتغيرات العشوائية في هذا الفصل على فضاء عينة واحدة . وفي هذا الفضاء فرضنا أن نتائج تجربة إحصائية هي قيم لمتغير عشوائى وحيد . ومع ذلك فإنه قد نكون مضطرين لاستخدام العديد من المتغيرات العشوائية في وقت واحد . فمثلا لو طلب إلينا في تجربة سحب ثلاث أوراق من ورق اللعب المكون من 52 ورقة حساب احتمال أن يكون بين الأوراق المسحوبة ورقة دينارى ورقتين كبة لكنا مضطرين إلى اعتبار المتغير  $X$  ممثلا لعدد أوراق الدينارى المسحوبة و  $Y$  لعدد أوراق الكبة المسحوبة في نفس التجربة .

وبشكل عام إذا كان  $Y, X$  متغيرين عشوائيين منقطعين ، وإذا رمزنا بـ  $f(x, y)$  لاحتمال أن يأخذ المتغير  $X$  القيمة  $x$  والمتغير  $Y$  القيمة  $y$  أى إن :

$$P [ X = x, Y = y ] = f(x, y)$$

فعندئذ تكون هذه الدالة ممثلة لاحتمال أن يفترض المتغير  $X$  القيمة  $x$  ، والمتغير  $Y$  القيمة  $y$  . ونطلق على هذه الدالة اسم دالة توزيع الاحتمال المشترك . فمثلا في المثال السابق الذى سقناه عند سحب ثلاث أوراق لعب يمثل  $f(1,2)$  الاحتمال المطلوب .

**تعريف (٢,٨) دالة توزيع الاحتمال المشترك Joint probability distribution function**

نسمى الدالة  $f(x,y)$  بدالة توزيع الاحتمال المشترك للمتغيرين العشوائيين المنقطعين  $X, Y$  إذا حققت الشروط الثلاثة التالية :

1.  $f(x,y) \geq 0$  for all  $(x,y)$
2.  $\sum_x \sum_y f(x,y) = 1$
3.  $P[(x,y) \in A] = \sum_A f(x,y)$  وذلك من أجل أى منطقة  $A$  فى المستوى  $xy$ .

لايجاد جميع قيم دالة توزيع الاحتمال المشترك لمتغيرين عشوائيين منقطعين  $Y, X$  قيمهما  $(x_1, \dots, x_n), (y_1, \dots, y_n)$  على الترتيب نسوق الجدول التالى :

الجدول (٢,٤) جدول التوزيع المشترك لمتغيرين منقطعين

$\begin{matrix} X \\ Y \end{matrix}$	$x_1$	$x_2$	$x_3$	$\dots$	$x_n$	
$y_1$	$f(x_1, y_1)$	$f(x_2, y_1)$	$f(x_3, y_1)$	$\dots$	$f(x_n, y_1)$	$\sum_{i=1}^n f(x_i, y_1)$
$y_2$	$f(x_1, y_2)$	$f(x_2, y_2)$	$f(x_3, y_2)$	$\dots$	$f(x_n, y_2)$	$\sum_{i=1}^n f(x_i, y_2)$
$y_3$	$f(x_1, y_3)$	$f(x_2, y_3)$	$f(x_3, y_3)$	$\dots$	$f(x_n, y_3)$	$\sum_{i=1}^n f(x_i, y_3)$
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
$y_n$	$f(x_1, y_n)$	$f(x_2, y_n)$	$f(x_3, y_n)$	$\dots$	$f(x_n, y_n)$	$\sum_{i=1}^n f(x_i, y_n)$
	$\sum_{i=1}^n f(x_i, y_i)$	$\sum_{i=1}^n f(x_i, y_i)$	$\sum_{i=1}^n f(x_i, y_i)$	$\dots$	$\sum_{i=1}^n f(x_i, y_i)$	

حيث يمثل العدد  $f(x_i, y_i)$  احتمال أن يفترض المتغير  $X$  القيمة  $x_i$  ، والمتغير  $Y$  القيمة  $y_i$  أى أن :

$$f(x_i, y_i) = P[X = x_i, Y = y_i]$$

## مثال (٢,١٠)

لنفرض في تجربة إلقاء حجرى نرد متماثلين ومتوازيين أن  $X$  يمثل القيمة العظمى للعددتين اللذين ظهرا، وأن  $Y$  يمثل مجموع العددين . نلاحظ أن فضاء العينة مؤلف من 36 نقطة عينة . انظر المثال (١,١٠) . أما قيم المتغيرين  $Y, X$  فهي على الترتيب :

$$X = \{ 1, 2, 3, 4, 5, 6 \}$$

$$Y = \{ 2, 3, 4, \dots, 12 \}$$

أما توزيع الاحتمال المشترك لهذين المتغيرين المنقطعين فيوضحه الجدول التالى :

جدول (٢,٥)

$Y \backslash X$	1	2	3	4	5	6	
2	1/36	0	0	0	0	0	1/36
3	0	2/36	0	0	0	0	2/36
4	0	1/36	2/36	0	0	0	3/36
5	0	0	2/36	2/36	0	0	4/36
6	0	0	1/36	2/36	2/36	0	5/36
7	0	0	0	2/36	2/36	2/36	6/36
8	0	0	0	1/36	2/36	2/36	5/36
9	0	0	0	0	2/36	2/36	4/36
10	0	0	0	0	1/36	2/36	3/36
11	0	0	0	0	0	2/36	3/36
12	0	0	0	0	0	1/36	1/36
	1/36	3/36	5/36	7/36	9/36	11/36	

والحقيقة أن قيم دالة توزيع الاحتمال المشترك  $f(x,y)$  قد تم حسابها بواسطة العلاقة :

$$f(x,y) = P [ X = x, Y = y ]$$

وعلى سبيل المثال نلاحظ أن :

$$f(2,4) = P [ X = x, Y = y ]$$

$$= P [ \text{مجموع الوجهين اللذين ظهرا 4 ، وأكبرهما 2} ]$$

$$= P [ (2,2) ]$$

$$= \frac{1}{36}$$

وهي نفس القيمة التي نلاحظها في الجدول (٢,٥) في العمود الثاني والسطر الرابع كذلك فإن :

$$f(3,4) = P [ X = 3, Y = 4 ]$$

$$= P [ \text{أكبر العددين اللذين ظهرا ثلاثة ، ومجموعهما أربعة} ]$$

$$= P [ (1,3) , (3,1) ] = \frac{2}{36}$$

وهي القيمة التي نجدها في العمود الثالث والسطر الرابع .

### تعريف (٢,٩) دالة الكثافة المشتركة لمتغيرين عشوائيين مستمرين

نسمى الدالة  $f(x,y)$  بدالة الكثافة المشتركة لمتغيرين عشوائيين مستمرين  $Y, X$  إذا حققت الشروط الثلاثة التالية :

$$1) f(x, y) \geq 0. \text{ for all } (x,y)$$

$$2) \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) dx dy = 1$$

$$3. \int_A \int f(x,y) dx dy = P[(X, Y) \in A]$$

وذلك مهما تكن المنطقة  $A$  من المستوى  $XY$ .

### مثال (٢,١١)

نفرض أن دالة الكثافة المشتركة للمتغيرين المستمرين  $(X,Y)$  هي من الشكل :

$$f(x,y) = \begin{cases} \frac{1}{8} (6-x-y) & : 0 < x < 2 \\ & 2 < y < 4 \\ 0 & : \text{فيما عدا ذلك} \end{cases}$$

بما أن دالة التوزيع التراكمي  $F$  لهذه الثنائية العشوائية تحدد بالعلاقة :

$$F(x,y) = \int_{u=-\infty}^x \int_{v=-\infty}^y f(u,v) du dv$$

فنعندئذ نجد أنه إذا كان  $x = 1$  ,  $y = 3$  فعندئذ يكون :

$$F(1,3) = \int_{x=0}^1 \int_{y=2}^3 \frac{1}{8} (6-x-y) dx dy =$$

$$F(1,3) = \frac{1}{8} \int_{x=0}^1 \left( 6y - xy - \frac{y^2}{2} \right)_{y=2}^{y=3} dx$$

$$= \frac{1}{8} \int_{x=0}^1 \left( \frac{7}{2} - x \right) dx = \frac{3}{8}$$

### ملاحظة

إذا علمت دالة توزيع الاحتمال المشتركة  $f(x,y)$  للمتغيرين  $X, Y$  فعندئذ يمكن معرفة توزيع الاحتمال لكل من  $Y, X$  كلا على حدة بواسطة العلاقتين التاليتين :

$$g(x) = \sum_y f(x,y)$$

$$h(y) = \sum_x f(x,y)$$

وذلك إذا كان المتغيران منقطعين .

أما إذا كان المتغيران مستمرين فإننا نستخدم العلاقتين التاليتين :



$$g(x) = \int_{-\infty}^{+\infty} f(x, y) dy$$

$$h(y) = \int_{-\infty}^{+\infty} f(x, y) dx$$

ونسَمي كلا من الدالتين  $h(y)$  و  $g(x)$  بالتوزيع الهامشي (marginal distribution) لكل من المتغيرين  $X, Y$  على الترتيب .

وللتأكد من أن  $g(x), h(y)$  يمثل كلا منهما توزيع احتمال ، علينا أن نتأكد بسهولة من تحقق شروط التعريف (٢,٤) ، أو التعريف (٢,٦) على التالى حسبما يكون المتغير منقطعا أو مستمرا على الترتيب . وبالحقيقة نلاحظ أن :

$$\int_{-\infty}^{+\infty} f(x) dx = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) dy dx = 1.$$

كذلك فإن :

$$P[a < X < b] = P[a < X < b, -\infty < Y < +\infty]$$

$$\begin{aligned} &= \int_{x=a}^b \int_{y=-\infty}^{+\infty} f(x, y) dy dx \\ &= \int_a^b g(x) dx \end{aligned}$$

### ملاحظة هامة

بالعودة إلى الجدول (٢,٤) الممثل للتوزيع المشترك لمتغيرين منقطعين ، نلاحظ أن العمود الأخير في هذا الجدول يمثل قيم التوزيع الهامشي للمتغير  $Y$  . فمثلا نلاحظ أن :

$$\begin{array}{lcl} y_1 & \xrightarrow{\text{يقابلها}} & \sum_{x_i} f(x_i, y_1) \\ y_2 & \xrightarrow{\text{يقابلها}} & \sum_{x_i} f(x_i, y_2) \end{array}$$

وهكذا ، كما نلاحظ أن عناصر السطر الأخير في نفس الجدول تمثل قيم التوزيع الهامشي للمتغير  $X$  ، وعلى سبيل المثال نلاحظ أن القيم  $x_1, x_2, x_3$  يقابلها على الترتيب القيم :

$$\sum_{j=1}^n f(x_1, y_j) , \quad \sum_{j=1}^n f(x_2, y_j) , \quad \sum_{j=1}^n f(x_3, y_j)$$

مثال (٢،١٢)

لنبحث عن قيم التوزيعات الهامشية لكل من  $Y, X$  في المثال (٢،١٠) ، نلاحظ أن التوزيع الاحتمالي للمتغير  $X$  في المثال (٢،١٠) يعطى بالجدول التالي :

X	1	2	3	4	5	6
P	$\frac{1}{36}$	$\frac{3}{36}$	$\frac{5}{36}$	$\frac{7}{36}$	$\frac{9}{36}$	$\frac{11}{36}$

كذلك نلاحظ أن توزيع الاحتمال بالنسبة للمتغير  $Y$  يمثل الجدول :

Y	2	3	4	5	6	7	8	9	10	11	12
P	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

مثال (٢،١٣)

لنبحث عن دالة الكثافة الاحتمالية لكل من المتغيرين المستمرين  $Y, X$  في المثال

(٢،١١)

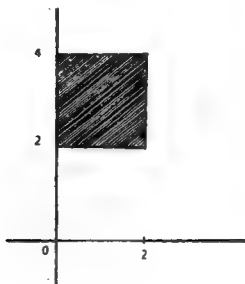
الحل

نلاحظ أن :

$$g(x) = \int_{-\infty}^{+\infty} f(x, y) dy$$

ومن عبارة الكثافة المشتركة في المثال (٢, ١١) نجد أن الكثافة الاحتمالية مجمعة في القسم المظلل من الشكل المرفق ، أما خارج المربع المظلل ، فإن الكثافة معدومة وعلى هذا فإن :

$$g(x) = \int_{y=2}^4 \frac{1}{8} (6 - x - y) dy$$



ومنه :

$$g(x) = \begin{cases} \frac{1}{8} (6 - 2x) & : 0 < x < 2 \\ 0 & : \text{فيما عدا ذلك} \end{cases}$$

كذلك فإن :

$$h(y) = \int_{-\infty}^{+\infty} f(x, y) dx = \int_{x=0}^2 \frac{1}{8} (6 - x - y) dx$$

$$h(y) = \begin{cases} \frac{5-y}{4} & : 2 < y < 4 \\ 0 & : \text{فيما عدا ذلك} \end{cases}$$

لنفرض أن  $X, Y$  متغيرين عشوائيين منقطعين معرفين على فضاء العينة  $S$  في تجربة معينة من الواضح أن المجموعة الجزئية  $(X = x)$  هي حادث في الفضاء  $S$  ، وكذلك الأمر بالنسبة لـ  $(Y = y)$  . فإذا رمزنا لهذين الحادثين بالرمزين :

$$A = [X = x], \quad B = [Y = y]$$

فإننا نستنتج باستخدام الاحتمال الشرطي للحادث  $B$  بمعلومية  $A$  الذي أوردنا تعريفه في (١,١١) أن :

$$P(B/A) = \frac{A \cap B}{P(A)} : P(A) > 0$$

ومنه :

$$P[Y = y | X = x] = \frac{P[X = x, Y = y]}{P[X = x]}$$

$$= \frac{f(x, y)}{g(x)} : g(x) > 0$$

ومن السهل التأكد من أن  $f(x, y) / g(x)$  يحقق كل شروط توزيع الاحتمال . فلو رمزنا للطرف الأيسر في العلاقة السابقة بالرمز  $f(y/x)$  لوجدنا أن :

$$f(y/x) = \frac{f(x, y)}{g(x)} : g(x) > 0$$

ونسى الدالة الجديدة  $f(y/x)$  بالتوزيع الشرطي للمتغير  $Y$  بمعلومية أن المتغير  $X$  قد أخذ القيمة  $x$  . ( هذا يفرض أن كلا المتغيرين  $X, Y$  منقطعان ) . وبشكل مماثل نعرف  $f(x/y)$  ليثل التوزيع الشرطي للمتغير  $X$  بمعلومية أن المتغير  $Y$  = بالعلاقة .

$$f(x/y) = \frac{f(x, y)}{h(y)} : h(y) > 0$$

نعرف الكثافة الاحتمالية الشرطية لمتغير عشوائى مستمر  $X$  بمعلومية أن  $Y = y$  بالعلاقة :

$$f(x/y) = \frac{f(x, y)}{h(y)} : h(y) > 0$$

أما بالنسبة للكثافة الاحتمالية الشرطية للمتغير المستمر  $Y$  بمعلومية أن  $X = x$  فنحدد بالعلاقة :

$$f(y/x) = \frac{f(x, y)}{g(x)} : g(x) > 0$$

إذا أردنا أن نبحث عن احتمال أن يقع المتغير العشوائي المستمر  $X$  في المجال  $(a, b)$  بمعلومية أن  $Y = y$  فعلياً أن نستخدم العلاقة :

$$P[a < X < b | Y = y] = \int_a^b f(x/y) dx$$

مثال (٢،١٤)

لنبحث في المثال (٢،١٠) عن كل من  $f(x/2)$  ،  $P[X = 1 | Y = 2]$

الحل

من الواضح أن :

$$f(x/2) = \frac{f(x, 2)}{h(2)}$$

ومن الجدول (٢،٥) نجد أن :

$$h(2) = \frac{1}{36}$$

$$f(x/2) = 36 f(x, 2) : x = 1, 2, 3, 4, 5, 6$$

لذلك فإن :

$$f(x/2) = \begin{cases} (36) f(1,2) = (36) \left( \frac{1}{36} \right) = 1 & : x = 1 \\ (36) f(2,2) = (36) (0) = 0 & : x = 2 \\ (36) f(6,2) = (36) (0) = 0 & : x = 6 \end{cases}$$

وهكذا نجد أن جدول توزيع المتغير العشوائي  $X$  هو :

$X$	1	2	3	4	5	6
$f(x/2)$	0	0	0	0	0	0

من الجدول نستنتج أن :

$$P[X = 1 | Y = 2] = 1$$

وهذه النتيجة تمثل احتمال أن يكون أكبر العددين اللذين ظهرا على وجهي حجري نرد عند إلقائها هو الواحد بمعلومية أن مجموع الوجهين يساوي 2 .

مثال (٢,١٥).

لنبحث في المثال (٢,١١) عن كل من :  $f(y/x)$  ,  $P[Y < 3 | X = \frac{1}{2}]$

الحل

وجدنا في المثال (٢,١٣) أن :

$$g(x) = \begin{cases} \frac{3-x}{4} & : 0 < x < 2 \\ 0 & : \text{فيما عدا ذلك} \end{cases}$$

ومن المعلوم أن :

$$f(y/x) = \frac{f(x,y)}{g(x)}$$

لذلك فإن :

$$F(y/x) = \begin{cases} \frac{(6-x-y)}{2(3-x)} & : 0 < x < 2 \\ & : 2 < y < 4 \\ 0 & : \text{فيما عدا ذلك} \end{cases}$$

ثم إن :

$$f(3/\frac{1}{2}) = \frac{1}{2} \cdot g(\frac{1}{2}) = \frac{5}{8}$$

كذلك فإن :

$$P[Y < 3 | X = \frac{1}{2}] = \frac{6}{10}$$

مثال (٢،١٦)

بفرض أن دالة الكثافة المشتركة للمتغيرين المستمرين  $X, Y$  هي من الشكل :

$$f(x, y) = \begin{cases} \frac{(1+5x^4)y^3}{8} & : 0 < x < 1 \\ & : 0 < y < 2 \\ 0 & : \text{فيما عدا ذلك} \end{cases}$$

فالمطلوب :

(١) التحقق من الشرط الثاني في التعريف (٢،٩) بالنسبة للدالة  $f(x, y)$

(٢) حساب الاحتمال  $P[(X, Y) \in A]$  حيث إن :

$$A = \{(x, y) | 0 < x < \frac{1}{2}, \frac{1}{4} < y < \frac{1}{2}\}$$

(٣) إيجاد كل من  $f(x/y), h(y), g(x)$

الحل

نلاحظ أن :

— أولاً —

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) dx dy = \int_{x=0}^1 \int_{y=0}^2 \frac{1}{8} (1 + 5x^4) y^3 dy dx$$

$$\int_{y=0}^2 \left( \frac{xy^3}{8} + \frac{x^5 y^3}{8} \right) \Big|_{x=0}^1 dy$$

$$= \int_0^2 \frac{1}{4} y^3 dy = 1$$

— ثانياً —

$$P[(X, Y) \in A] = \int_{x=0}^1 \int_{y=\frac{1}{4}}^{\frac{1}{2}} \frac{(1 + 5x^4) y^3}{8} dy dx$$

$$= \frac{(15)(17)}{(16)^2 (32)^2} = 0.00097$$

— ثالثاً —

$$g(x) = \int_{y=-\infty}^{+\infty} f(x, y) dy = \int_{y=0}^2 \frac{1}{8} (1 + 5x^4) y^3 dy$$

ومنه :

$$g(x) = \begin{cases} \frac{1}{2} (1 + 5x^4) & : 0 < x < 1 \\ 0 & : \text{فيما عدا ذلك} \end{cases}$$

ثم أن :

$$h(y) = \int_{x=-\infty}^{+\infty} f(x, y) dx = \int_{x=0}^1 \frac{1}{8} (1 + 5x^4) y^3 dx$$



$$h(y) = \begin{cases} \frac{y^3}{4} & : 0 < y < 2 \\ 0 & : \text{ما عدا ذلك} \end{cases}$$

وأخيراً فإن :

$$f(x/y) = \frac{f(x,y)}{h(y)} \\ = \frac{(1+5x^4)}{2}$$

ومن ثم يكون :

$$f(x/y) = \begin{cases} \frac{(1+5x^4)}{2} & : 0 < x < 1, 0 < y < 2 \\ 0 & : \text{فيما عدا ذلك} \end{cases}$$

والملاحظ في هذا التمرين أن  $f(x/y) = g(x)$  ، أى إن  $f(x,y) = g(x) h(y)$  ، وهذا يقودنا إلى التعريف التالى :

**تعريف (٢,١٠) المتغيران العشوائيان المستقلان**

**Two independence the random variables**

بفرض أن  $Y, X$  متغيرين عشوائيين منقطعان أو مستمران بالكثافة الاحتمالية المشتركة  $f(x,y)$  ، والكثافات الهامشية  $g(x)$  ،  $h(y)$  على الترتيب ، عندئذ نقول بأن المتغيرين العشوائيين  $Y, X$  مستقلان إحصائياً إذا وفقط إذا كان :

$$f(x,y) = g(x) \cdot h(y)$$

من أجل جميع قيم  $(x,y)$  نلاحظ أن المتغيرين العشوائيين المستمرين  $Y, X$  الواردين في المثال (٢,١٦) السابق مستقلان إحصائياً لأنهما يحققان التعريف (٢,١٠) .

## مثال (٢, ١٧)

$X, Y$  متغيران عشوائيان منقطعان لهما توزيع احتمالي مشترك معطى بالجدول

التالي :

$X \backslash Y$	4	10	
1	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{2}$
3	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{2}$
	$\frac{1}{2}$	$\frac{1}{2}$	

لنبحث فيما إذا كان هذان المتغيران مستقلين إحصائياً أم لا . من الجدول السابق نلاحظ أن جدول توزيع المتغير  $X$  هو :

$X$	4	10
$g(x)$	$\frac{1}{2}$	$\frac{1}{2}$

وأن جدول توزيع  $Y$  هو :

$Y$	1	3
$h(x)$	$\frac{1}{2}$	$\frac{1}{2}$

كما نلاحظ أنه من أجل جميع قيم  $(X, Y)$  فإن :

$$f(x, y) = g(x) \cdot h(y)$$

وهذا ما يشير إلى أن المتغيرين العشوائيين  $X, Y$  مستقلان إحصائياً .

## مثال (٢, ١٨)

بفرض أن التوزيع الاحتمالي المشترك للمتغيرين العشوائيين  $X, Y$  معطى بالجدول التالي :

$X \backslash Y$	4	10	
1	0	$\frac{1}{2}$	$\frac{1}{2}$
3	$\frac{1}{2}$	0	$\frac{1}{2}$
	$\frac{1}{2}$	$\frac{1}{2}$	

لنبحث عن جدول توزيع كل متغير من هذين المتغيرين ، ثم لندرس استقلال هذين المتغيرين .

## الحل

من الواضح من جدول التوزيع المشترك أن قيم المتغير  $X$  هي 4, 10 وأن احتمالات هذه القيم هي  $\frac{1}{2}$  ،  $\frac{1}{2}$  على الترتيب ، وعلى هذا الأساس فإن جدول توزيع المتغير الأول هو :

$X$	4	10
$g(x)$	$\frac{1}{2}$	$\frac{1}{2}$

كما نلاحظ بنفس الطريقة أن جدول توزيع  $Y$  هو :

$Y$	1	3
$h(x)$	$\frac{1}{2}$	$\frac{1}{2}$

ولو شكلنا الجداء  $h(y) \cdot g(x)$  ، وقارناه مع قيم  $f(x,y)$  لجميع القيم  $(x,y)$  لوجدنا أن  $h(y) \cdot g(x) = f(x,y)$  ، وهذا يعنى أن المتغيرين  $X, Y$  ليسا مستقلين .

يمكن تعميم جميع التعاريف السابقة والمتعلقة بمتغيرين عشوائيين إلى حالة  $n$  متغير .  
فبفرض أن  $f(x_1, \dots, x_n)$  يمثل التوزيع الاحتمالى المشترك للمتغيرات العشوائية  $X_1, \dots, X_n$  ، فإننا نلاحظ أن التوزيع الهامشى للمتغير  $X_1$  مثلاً يعطى بالعلاقة :

$$g(x_1) = \sum_{x_2} \sum_{x_3} \dots \sum_{x_n} f(x_1, \dots, x_n)$$

وذلك عندما تكون المتغيرات منقطعة .  
وبالعلاقة :

$$g(x_1) = \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} f(x_1, \dots, x_n) dx_2 dx_3 \dots dx_n$$

في الحالة التى تكون فيها المتغيرات السابقة مستمرة . كما نلاحظ أن التوزيع الهامشى المشترك والذي نرمز له بالرمز  $\phi(x_1, x_2)$  يمكن إيجاداه بإحدى العلاقتين :

(١) إذا كانت المتغيرات منقطعة :

$$\phi(x_1, x_2) = \sum_{x_3} \dots \sum_{x_n} f(x_1, \dots, x_n)$$

(٢) إذا كانت المتغيرات مستمرة :

$$= \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} f(x_1, \dots, x_n) dx_3 dx_4 \dots dx_n$$

نلاحظ أيضاً أن التوزيع الشرطى المشترك joint cinditional distribution للمتغيرات  $X_1, X_2, X_3$  معلومية أن  $X_n = x_n, \dots, X_5 = x_5, X_4 = x_4$  يعطى بالعلاقة :

$$f(x_1, x_2, x_3 / x_4, x_5, \dots, x_n) = \frac{f(x_1, \dots, x_n)}{g(x_4, \dots, x_n)}$$

حيث يمثل  $(x_1, \dots, x_n)$  التوزيع الهامشي المشترك للمتغيرات العشوائية  $X_1, X_2, \dots, X_n$ .

أما تعريف الاستقلال الإحصائي لجملة متغيرات عشوائية فسنوفه من خلال التعريف التالي :

### تعريف (٢،١١) المتغير العشوائي المستقل

#### The independent random variable

ليكن  $(X_1, X_2, \dots, X_n)$  متغيراً عشوائياً منقطعاً أو مستمراً بالتوزيع الاحتمالي المشترك  $f(x_1, x_2, \dots, x_n)$  والتوزيعات الهامشية  $f_1(x_1), f_2(x_2), \dots, f_n(x_n)$  على التوالي :  
نقول بأن المتغيرات  $X_1, X_2, \dots, X_n$  مستقلة إحصائياً وبالتبادل إذا وفقط إذا تحققت العلاقة :

$$f(x_1, \dots, x_n) = f_1(x_1) f_2(x_2) \dots f_n(x_n)$$

### مثال (٢،١٩)

نفرض أن  $X_1, X_2, X_3$  تمثل ثلاثة متغيرات عشوائية مستقلة إحصائياً وبالتبادل ، ولكل منها كثافة احتمالية معطاة بالدالة :

$$f(x) = \begin{cases} e^{-x} & : x > 0 \\ 0 & : \text{فيما عدا ذلك} \end{cases}$$

لنحسب الاحتمال :

$$P[X_1 < 3, 1 < X_2 < 2, X_3 > 2]$$

### الحل

من الاستقلال التبادلي للمتغيرات  $X_1, X_2, X_3$  وحسب التعريف (٢،١١)

نجد أن :  $f(x_1, x_2, x_3) = f(x_1) \cdot f(x_2) \cdot f(x_3)$

$$= e^{-x_1} e^{-x_2} e^{-x_3}$$

$$= e^{-(x_1+x_2+x_3)} : x_1 > 0, x_2 > 0, x_3 > 0$$

وهكذا فإن :

$$P[X_1 < 3, 1 < X_2 < 2, X_3 > 2] =$$

$$\int_{x_1=0}^3 \int_{x_2=0}^2 \int_{x_3=0}^2 e^{-(x_1+x_2+x_3)} dx$$

$$(1 - e^{-3}) (e^{-1} - e^{-2}) (1 - e^{-2}) = 0.191$$

مثال (٢,٢٠)

X, Y, Z ثلاثة متغيرات عشوائية مستمرة دالة كثافتها المشتركة من الشكل :

$$f(x, y, z) = \begin{cases} \frac{4xyz^2}{9} & : 0 < x < 1, 0 < y < 1, 0 < z < 3 \\ 0 & : \text{فيما عدا ذلك} \end{cases}$$

نلاحظ أن الكثافة المشتركة للمتغيرين (Y, Z) هي من الشكل :

$$\phi(y, z) = \int_{x=-\infty}^{+\infty} f(x, y, z) dx = \int_{x=0}^1 \frac{4}{9} xyz^2 dx$$

ومنه :

$$\phi(y, z) = \begin{cases} \frac{2}{9} yx^2 & : 0 < y < 1, 0 < z < 3 \\ 0 & : \text{فيما عدا ذلك} \end{cases}$$

أما الكثافة الهامشية للمتغير  $Y$  فيمكن إيجادها من الدالة السابقة على النحو التالي :

$$\begin{aligned} h(y) &= \int_{-\infty}^{+\infty} \phi(y, z) dz = \int_{z=0}^3 \frac{2}{9} yz^2 dz \\ &= \frac{2}{9} y \left( \frac{z^3}{3} \right)_{z=0}^3 \end{aligned}$$

وهكذا نجد أن :

$$h(y) = \begin{cases} 2y & : 0 < y < 1 \\ 0 & : \text{فيما عدا ذلك} \end{cases}$$

كذلك فإن الكثافة الشرطية للمتغير  $X$  بعلومية أن  $Z = 2$  ،  $Y = \frac{1}{2}$  هي

$$f(x / \frac{1}{2}, 2) = \frac{f(x, \frac{1}{2}, 2)}{\phi(\frac{1}{2}, 2)} = \begin{cases} 2x & : 0 < x < 1 \\ 0 & : \text{فيما عدا ذلك} \end{cases}$$

وأخيرا فإن :

$$f(x / \frac{1}{2}, 2) = \begin{cases} 2x & : 0 < x < 1 \\ 0 & : \text{فيما عدا ذلك} \end{cases}$$

### (٢,٦) التوقع الرياضي The mathematical expectation

لنلقى قطعتى نقود عشر مرات متتاليات ، ولنرمز بـ  $X$  لعدد مرات الصورة التي ظهرت في كل إلقاء . نلاحظ أن قيم المتغير  $X$  هي 0, 1, 2 . لنفرض أنه خلال التجربة لم تظهر الصورة ثلاث مرات على التوالي ، وظهرت صورة واحدة خمس مرات ، وأخيرا ظهرت صورتين ، نلاحظ أن متوسط عدد مرات الصورة التي ظهرت في كل إلقاء هو :

$$\frac{(0)(3) + (1)(5) + (2)(2)}{10} = (0) \left( \frac{3}{10} \right) + (1) \left( \frac{5}{10} \right) + (2) \left( \frac{2}{10} \right) = 0.9$$

هذا العدد يمثل قيمة وسطى وليس من الضروري أن يكون نتيجة ممكنة للتجربة . كما نلاحظ أن الكسور  $\frac{3}{10}, \frac{5}{10}, \frac{2}{10}$  قد ضربت على التالى بالأعداد 0, 1, 2 وهذه الكسور تمثل التكرار النسبي للنتائج المختلفة .

لنحسب على المدى الطويل متوسط عدد مرات الصورة في كل قذفة ، ولنرمز لهذا المتوسط والذي نسميه بالتوقع الرياضى أو القيمة المتوقعة بالرمز  $E(X)$  . من تعريف التكرار النسبى أو الاحتمال نستطيع أن نتوقع على المدى البعيد علام ظهور صورة في حوالى ربع الوقت ، وظهور صورة واحدة في نصف الوقت ، وظهور صورتين في ربع الوقت الباقى لذلك فإن :

$$E(X) = (0) \left( \frac{1}{4} \right) + (1) \left( \frac{1}{2} \right) + (2) \left( \frac{1}{4} \right) = 1$$

ونتوقع أنه ظهور صورة في المتوسط في كل مرة عندما نلقى قطعتى نقود مرة تلو الأخرى .

من المناقشة السابقة نستنتج أن التوقع الرياضى أو المتوسط يمكن حسابه بضرب كل قيمة من قيم المتغير العشوائى بالاحتمال الموافق ، ثم يجمع هذه الجداءات . وهذا صحيح بالطبع إذا كان المتغير منقطعا . أما إذا كان المتغير مستمرا فَنستبدل الجمع بالتكامل .



**تعريف (٢,١٢) التوزيع الاحتمالي Probability distribution**

بفرض أن  $f(x)$  دالة الكثافة (أو التوزيع الاحتمالي) للمتغير العشوائي  $X$ . عندئذ نعرف التوقع الرياضي لهذا المتغير بإحدى العالقتين :

$$E(X) \approx \begin{cases} \sum_x x f(x) & \text{إذا كان المتغير منقطعاً :} \\ \int_{-\infty}^{+\infty} x f(x) dx & \text{إذا كان المتغير مستمراً :} \end{cases}$$

يشرط لوجود التوقع الرياضي تحقق  $\sum |x| f(x) < \infty$  أو  $\int_{-\infty}^{+\infty} |x| f(x) dx < \infty$

مثال (٢,٢١)

لنحسب التوقع الرياضي لمتغير عشوائي منقطع  $X$  جدول توزيع من الشكل :

$X$	-1	0	1	2
$f(x)$	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{3}{8}$	$\frac{1}{4}$

نلاحظ أن توقع هذا المتغير هو :

$$E(X) = (-1) \left( \frac{1}{8} \right) + (0) \left( \frac{1}{4} \right) + (1) \left( \frac{3}{8} \right) + (2) \left( \frac{1}{4} \right)$$

ومنه :

$$E(X) = \frac{3}{4}$$

مثال (٢,٢٢)

لنبحث عن التوقع الرياضي لمتغير عشوائي  $Y$  دالة كثافته من الشكل :

$$f(y) = \frac{1}{\sqrt{2\pi}} e^{-1/2 y^2} : -\infty < x < +\infty$$

نلاحظ أن التوقع الرياضي لهذا المتغير هو :

$$\begin{aligned} E(Y) &= \int_{-\infty}^{+\infty} y \cdot f(y) dy = \int_{-\infty}^{+\infty} y \cdot \frac{1}{\sqrt{2\pi}} e^{-1/2 y^2} dy \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} d \left( e^{-1/2 y^2} \right) \\ &= \frac{1}{\sqrt{2\pi}} \left( e^{-1/2 y^2} \right)_{-\infty}^{+\infty} = 0 \end{aligned}$$

مثال (٢,٢٣)

نلاحظ أن التوقع الرياضي للمتغير  $X$  في المثال (٢,٧) يساوى :

$$\begin{aligned} E(X) &= \int_{-\infty}^{+\infty} x \cdot f(x) dx = \int_1^3 x \cdot \frac{1}{2} dx = \frac{1}{4} \left( x^2 \right)_{x=1}^3 \\ &= \frac{1}{4} (9 - 1) = 2 \end{aligned}$$

مثال (٢,٢٤)

$X$  متغير عشوائى دالة كثافته الاحتمالية من الشكل :

$$f(x) = \begin{cases} \frac{4}{\pi(1+x^2)} & : 0 < x < 1 \\ 0 & : \text{فيما عدا ذلك} \end{cases}$$

لنحسب التوقع الرياضى لهذا المتغير . نلاحظ أن :

$$\begin{aligned}
 E(X) &= \int_{-\infty}^{\infty} x \cdot f(x) dx = \int_0^1 \frac{4x}{\pi(1+x^2)} dx \\
 &= \frac{2}{\pi} \int_0^1 \frac{d(1+x^2)}{(1+x^2)} = \frac{2}{\pi} [\ln(1+x^2)]_0^1 \\
 &= \frac{2}{\pi} [\ln 2 - \ln 1] \\
 &= \frac{\ln 4}{\pi}
 \end{aligned}$$

مثال (٢,٢٥)

لنبحث عن التوقع الرياضى للمتغير العشوائى  $X$  الممثل لعمر نوع معين من اللببات بالساعات ، إذا فرضنا أن الكثافة الاحتمالية لهذا العمر معطاة بالعلاقة :

$$f(x) = \begin{cases} \frac{375000}{x^4} & : x > 50 \\ 0 & : \text{فيما عدا ذلك} \end{cases}$$

نلاحظ أن التوقع الرياضى لعمر هذا النوع من اللببات هو :

$$E(X) = \int_{50}^{\infty} x \cdot \frac{375000}{x^4} dx = 75$$

ولهذا فإننا نستطيع أن نتوقع أن نخدم كل لبة بمعدل 75 ساعة . لنفرض أن  $g(x)$  تمثل دالة ما بالمتغير العشوائى  $X$  ذى الكثافة  $f(x)$  مثلاً :

$$g(x) = \begin{cases} 2x - 1 \\ x^3 \\ \sqrt{x} \\ x^2 \end{cases}$$

من الملاحظ أنه إذا أخذ المتغير  $X$  القيم  $-1, 0, 1, 2$  ، وكان  $g(x) = x^2$  فعندئذ ستكون قيم المتغير الجديد  $g(x)$  هي  $0, 1, 2$  ، أما الاحتمالات الموافقة لهذه القيم فنحسبها على النحو التالي :

$$P [ g(x) = 0 ] = P (X = 0) = f(0)$$

$$P [ g(x) = 1 ] = P [ (X = -1) ] + P [ (X = 1) ] = f(-1) + f(1)$$

$$P [ g(x) = 4 ] = P(X = 2) = f(2)$$

وسيكون جدول قانون توزيع  $g(x)$  من الشكل :

y	0	1	4
$P[g(x) = y]$	$f(0)$	$f(-1) + f(1)$	$f(2)$

ويكون التوقع الرياضي لهذا المتغير في هذه الحالة من الشكل :

$$E [ g(x) ] = \sum_y y \cdot P [ g(x) = y ]$$

$$= 0 \cdot P [ g(x) = 0 ] + 1 \cdot P [ g(x) = 1 ] + 4 \cdot P [ g(x) = 4 ]$$

$$= 0 \cdot f(0) + 1 [ f(-1) + f(1) ] + 4 f(2)$$

$$= \sum_y g(x) \cdot f(x)$$

يمكن تعميم النتيجة السابقة بالنظرية (٢,١) من أجل المتغيرات المستمرة ، والمنقطعة .

### نظرية (٢,١)

بفرض  $X$  متغيراً عشوائياً بالكثافة الاحتمالية  $f(x)$  ، ولتكن  $g(x)$  دالة ما بهذا المتغير . عندئذ يحسب التوقع الرياضي لهذه الدالة بإحدى العلاقتين التاليتين :

$$E [ g(x) ] = \sum_x g(x) \cdot f(x)$$

إذا كان  $X$  منقطعاً :

إذا كان  $X$  مستمراً :

$$= \int_{x=-\infty}^{+\infty} g(x) f(x) dx$$

مثال (٢,٢٦)

بفرض  $X$  هو المتغير المعرف بالمثال (٢,٢١) ، وإن  $g(X) = X^2$  . ولنحسب التوقع الرياضي للمتغير الجديد  $g(X)$  .

نلاحظ في المثال (٢,٢١) أن المتغير  $X$  هو متغير منقطع ، ولهذا فإن للمتغير الجديد  $g(X)$  جدول توزيع من الشكل :

$X$	1	0	4
$f_{X^2}(y)$	$\frac{4}{8}$	$\frac{1}{4}$	$\frac{1}{4}$

ويكون التوقع الرياضي لـ  $g(X)$  مساوياً :

$$E[g(x)] = 1 \cdot \frac{4}{8} + 0 \cdot \frac{1}{4} + 4 \cdot \frac{1}{4}$$

$$= 1.5$$

مثال (٢,٢٧)

لنحسب في المثال (٢,١٨) كلا من  $E(X)$  و  $E(X^2)$  نعلم أن :

$X$	4	10
$f(x)$	$\frac{1}{2}$	$\frac{1}{2}$

ولذلك فإن :

$$E(x) = 4 \cdot \frac{1}{2} + 10 \cdot \frac{1}{2} = 7$$

كذلك فإن :

$X^2$	16	100
$f(x^2)$	$\frac{1}{2}$	$\frac{1}{2}$

من هذا الجدول نستنتج أن :

$$E(X^2) = (16) \left( \frac{1}{2} \right) + (100) \left( \frac{1}{2} \right) = 58$$

مثال (٢,٢٨)

نلاحظ في المثال (٢,١٠) أن للمتغير العشوائي المنقطع  $X$  والممثل للقيمة العظمى لوجهي حجري نرد مقنوفين ، جدول توزيع من الشكل :

$X$	1	2	3	4	5	6
$f(x)$	$\frac{1}{36}$	$\frac{3}{36}$	$\frac{5}{36}$	$\frac{7}{36}$	$\frac{9}{36}$	$\frac{11}{36}$

لنحسب توقع هذا المتغير ، ثم توقع مربعة نلاحظ أولاً أن :

$$E(X) = \sum_x x \cdot f(x) = 1 \cdot \left( \frac{1}{36} \right) + 2 \cdot \left( \frac{3}{36} \right) + 3 \cdot \left( \frac{5}{36} \right) + 4 \cdot \left( \frac{7}{36} \right) +$$

$$5 \cdot \left( \frac{9}{36} \right) + 6 \cdot \left( \frac{11}{36} \right)$$

ومنه :

$$E(X) = 4.472$$

كما أن جدول توزيع المتغير  $X^2$  هو :

$X^2$	1	4	9	16	25	36
$f(x^2)$	$\frac{1}{36}$	$\frac{3}{36}$	$\frac{5}{36}$	$\frac{7}{36}$	$\frac{9}{36}$	$\frac{11}{36}$

والتوقع الرياضى لهذا المتغير الأخير هو :

$$\begin{aligned}
 E(X^2) &= \sum_{x^2} x^2 \cdot f(x) \\
 &= 1 \cdot \left(\frac{1}{36}\right) + 4 \cdot \left(\frac{3}{36}\right) + 9 \cdot \left(\frac{5}{36}\right) + 16 \cdot \left(\frac{7}{36}\right) + \\
 &\quad 25 \cdot \left(\frac{9}{36}\right) + 36 \cdot \left(\frac{11}{36}\right) \\
 &= 21.972
 \end{aligned}$$

مثال (٢,٢٩)

متغير عشوائى دالة كثافته الاحتمالية من الشكل :

$$f(x) = \begin{cases} \frac{x^2}{3} & : -1 < x < 2 \\ 0 & : \text{فيما عدا ذلك} \end{cases}$$

لنبحث عن التوقع الرياضى للمتغير  $g(x) = 5x - 1$   
نلاحظ أن :

$$E[g(x)] = \int_{-1}^2 (5x - 1) \cdot \frac{x^2}{3} dx$$

$$E[5x - 1] = \frac{1}{3} \int_{-1}^2 (5x^3 - x^2) dx = \frac{1}{3} \left[ \frac{5}{4} x^4 - \frac{1}{3} x^3 \right]_{-1}^2 = \frac{21}{4} \quad \text{ومنه :}$$

سنعمم تعريف التوقع الرياضى إلى حالة متغيرين عشوائيين  $X, Y$  بالكثافة الاحتمالية المشتركة  $f(x, y)$

### تعريف (٢، ١٣) التوقع الرياضى The mathematical expectation

بفرض أن  $X, Y$  متغيران عشوائيان بالكثافة الاحتمالية المشتركة  $f(x, y)$ . نعرف التوقع الرياضى للدالة  $g(X, Y)$  بأحدى العلاقاتين :

(١) إذا كان المتغيران منقطعين

$$E[g(X, Y)] = \sum_x \sum_y g(x, y) f(x, y)$$

(٢) إذا كان المتغيران مستمرين

$$= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x, y) f(x, y) dx dy$$

إن تعميم التعريف (٢، ١٣) لحساب التوقع الرياضى لعدة متغيرات عشوائية ينتج مباشرة من التعريف نفسه :

### مثال (٢، ٣٠)

لنحسب في المثال (٢، ١٧) التوقع الرياضى للدالة  $g(X, Y) = X \cdot Y$

من التعريف (٢، ١٣) نعلم أن :

$$E(X \cdot Y) = \sum_x \sum_y x \cdot y \cdot f(x, y)$$

وبالتعويض نجد أن :

$$= (4) \cdot (1) \cdot \left(\frac{1}{4}\right) + (4) \cdot (3) \cdot \left(\frac{1}{4}\right) + (10) \cdot (1) \cdot \left(\frac{1}{4}\right) + (10) \cdot (3) \cdot \left(\frac{1}{4}\right)$$

$$= 1 + 3 + \frac{5}{2} + \frac{15}{2} = 14$$

### مثال (٢، ٣١)

لنبحث في المثال (٢، ١٦) عن التوقع الرياضى للدالة  $g(X, Y) = X^2 \cdot Y$



الحل :

نلاحظ أن المتغيرين العشوائيين الواردين في المثال (٢،١٦) مستمران ، لذلك فإن التوقع الرياضي للدالة  $g(X, Y) = X^2 \cdot Y$  نحسب بواسطة التعريف (٢،١٣) .

$$\begin{aligned} E(X^2 \cdot Y) &= \int_{y=0}^2 \int_{x=0}^1 x^2 \cdot y \frac{(1+5x^4)y^3}{8} dx dy \\ &= \frac{1}{8} \int_{y=0}^2 \int_{x=0}^1 (x^2y^4 + 5y^4x^6) dx dy \\ &= \int_{y=0}^2 \frac{11}{84} y^4 dy \\ &= \frac{88}{105} \end{aligned}$$

مثال (٢،٢٢)

بفرض أن  $X, Y$  متغيران عشوائيان مستمران معرفان بالمثال (٢،١٦) ، ولنبحث عن  $E(Y \cdot X = x)$  .

الحل :

من الواضح أن :

$$\begin{aligned} E(Y|X) &= \int_{y=0}^2 y \cdot f(y/x) dy \\ &= \int_{y=0}^2 y \cdot \frac{y^3}{4} dy \\ &= \frac{1}{4} \int_{y=0}^2 y^4 dy = \frac{8}{5} \end{aligned}$$

ملاحظة

بفرض أن  $g(X, Y) = X$  ، وبالتعويض في العلاقات الواردة في التعريف (٢،١٣) فإننا نحصل على العلاقتين العامتتين التاليتين :

(١) ( في حالة متغيرات منقطعة ) :

$$E(X) = \sum_x \sum_y x \cdot f(x,y) = \sum_x x \cdot g(x)$$

(٢) ( في حالة متغيرات مستمرة ) :

$$= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x \cdot f(x, y) \, dx \, dy = \int_{-\infty}^{+\infty} x \cdot g(x) \, dx$$

حيث يمثل  $g(x)$  التوزيع الهامشي للمتغير  $X$  . من هذا نستنتج أنه لحساب التوقع الرياضى للمتغير  $X$  في المستوى فإن بإمكاننا استخدام التوزيع المشترك للمتغيرين  $X, Y$  ، أو التوزيع الهامشي للمتغير  $X$  . وبنفس الطريقة نعرف :

(١) ( في حالة متغيرات منقطعة ) :

$$E(Y) = \sum_x \sum_y y \cdot f(x,y) = \sum_y y \cdot h(y)$$

(٢) ( في حالة متغيرات مستمرة ) :

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} y \cdot f(x, y) \, dx \, dy = \int_{-\infty}^{+\infty} y \cdot h(y) \, dy$$

حيث يمثل  $h(y)$  التوزيع الهامشي للمتغير  $Y$  .

### (٢،٧) قوانين التوقع الرياضى Laws of Expectation

سنبرهن الآن بعض القوانين المفيدة في حساب التوقع الرياضى ، وهذه الدساتير أو النظريات ستسمح لنا بحساب التوقعات الرياضية بوساطة بعض التوقعات المعروفة ، وستكون جميع النتائج صحيحة من أجل المتغيرات المنقطعة والمستمرة على حد سواء ، وسنقوم ببرهان هذه النتائج من أجل متغيرات مستمرة .

### نظرية (٢,٢)

إذا كان  $a$  ثابتين فإن :

$$E(aX + b) = a \cdot E(X) + b$$

البرهان :

من تعريف التوقع الرياضي نجد أن :

$$E(aX + b) = \int_{-\infty}^{+\infty} (ax + b) \cdot f(x) dx$$

$$= a \cdot \int_{-\infty}^{+\infty} x \cdot f(x) dx + b \cdot \int_{-\infty}^{+\infty} f(x) dx$$

والتكامل الثاني في الطرف الأيمن ما هو إلا إجراء ضرب  $a$  في التوقع الرياضي لـ  $X$  ، أما التكامل الأول فيساوى الواحد . لذلك فإن :

$$E(aX + b) = a \cdot E(x) + b$$

### نتيجة (٢,١)

بفرض  $a = 0$  ، فإننا نلاحظ أن  $E(b) = b$  أى إن التوقع الرياضى لثابت يساوى الثابت نفسه .

### نتيجة (٢,٢)

بفرض  $b = 0$  ، فإننا نلاحظ أن  $E(a \cdot X) = a \cdot E(X)$  أى إن التوقع الرياضى لجراء ضرب ثابت بمتغير عشوائى يساوى إلى حاصل ضرب الثابت بالتوقع الرياضى لذلك المتغير .

## نظرية (٢,٣)

إن التوقع الرياضى لمجموع أو الفرق بين دالتين ( بالمتغير العشوائى  $X$  ) يساوى مجموع أو الفرق بين فضل التوقع الرياضى لهاتين الدالتين :

$$E [ g(X) \pm h(X) ] = E [ g(X) ] \pm E [ h(X) ]$$

## البرهان

من تعريف التوقع الرياضى نجد أن :

$$\begin{aligned} E [ g(X) \pm h(X) ] &= \int_{-\infty}^{+\infty} [ g(x) \pm h(x) ] \cdot f(x) dx \\ &= \int_{-\infty}^{+\infty} g(x) \cdot h(x) dx \pm \int_{-\infty}^{+\infty} h(x) f(x) dx \\ &= E [ g(X) ] \pm E [ h(X) ] \end{aligned}$$

## مثال (٢,٣٣)

بفرض  $X$  هو المتغير العشوائى الوارد فى المثال (٢,١٦) ، احسب التوقع الرياضى للدالة  $(X + 1)^2$  .

## الحل

نعلم أن دالة الكثافة  $X$  فى المثال (٢,١٦) هى من الشكل :

$$g(x) = \begin{cases} \frac{1}{2} (1 + 5x^4) & : 0 < x < 1 \\ 0 & : \text{فيما عدا ذلك} \end{cases}$$

ومن الواضح أن :

$$\begin{aligned} E(X+1)^2 &= E(X^2 + 2X + 1) \\ &= E(X^2) + 2E(X) + 1 \end{aligned}$$

كما نجد :

$$\begin{aligned} E(X) &= \int_0^1 x \cdot \frac{1}{2} (1 + 5x^4) dx = \frac{2}{3} \\ E(X^2) &= \int_0^1 x^2 \cdot \frac{1}{2} (1 + 5x^4) dx = \frac{11}{21} \end{aligned}$$

وبالتعويض نجد أن :

$$\begin{aligned} E(X+1)^2 &= \frac{11}{21} + 2 \cdot \frac{2}{3} + 1 \\ &= \frac{60}{21} \end{aligned}$$

مثال (٢,٣٤)

وجدنا في المثال (٢,٣١) أن  $E(X) = \frac{3}{4}$  ، لنفتش عن  $E(2X+1)$  من الواضح أن :

$$E(2X+1) = 2 \cdot E(X) + 1$$

وبالتعويض نجد أن :

$$= 2 \cdot \frac{3}{4} + 1 = \frac{5}{2}$$

نفرض الآن أن لدينا متغيرين عشوائيين  $X, Y$  بالكثافة الاحتمالية المشتركة  $f(x,y)$  ، ولندرس نظرية تتعلق بمجموع وفرق دالتين بهذين المتغيرين .

نظرية (٢,٤)

إن التوقع الرياضي لمجموع أو فرق دالتين أو أكثر بهذين المتغيرين يساوى إلى مجموع أو فرق التوقعين الرياضيين لهاتين الدالتين .

$$E[g(X, Y) \pm h(X, Y)] = E[g(X, Y)] \pm E[h(X, Y)]$$

## البرهان

من التعريف (٢،١٣) نجد أن :

$$\begin{aligned} E [ g(X, Y) \pm h(X, Y) ] &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} [ g(x,y) \pm h(x,y) ] f(x,y) dx dy \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} h(x,y) \cdot f(x,y) dx dy \quad \pm \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x,y) f(x,y) dx dy \\ &= E [ g(X, Y) ] \pm E [ h(X, Y) ] \end{aligned}$$

## نتيجة (٢،٣)

بفرض أن  $h(X, Y) = Y, g(X, Y) = X$  فإننا نجد أن :

$$E(X \pm Y) = E(X) \pm E(Y)$$

وهذه النتيجة الأخيرة هامة جدا ، فهي تقرر أن التوقع الرياضى لمجموع ( أو فرق ) متغيرين عشوائيين ماهو إلا مجموع ( أو فرق ) التوقعين الرياضيين لهذين المتغيرين على الترتيب .

## نظرية (٢،٥)

إذا كان المتغيران العشوائيان  $X, Y$  مستقلين ، فإن :

$$E(X \cdot Y) = E(X) \cdot E(Y)$$

## البرهان

من تعريف التوقع الرياضى نجد أن :

$$E(X \cdot Y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x \cdot y \cdot f(x, y) dx \cdot dy$$

وبما أن  $X, Y$  مستقلان فيمكننا أن نكتب :

$$f(x, y) = g(x) \cdot h(y)$$

حيث يمثل كلا من  $g(x), h(y)$  التوزيعات الهامشية لكل من  $Y, X$  على الترتيب ، ولذلك فإن :

$$\begin{aligned}
 E(X, Y) &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x \cdot y \cdot g(x) \cdot h(y) \, dx \cdot dy \\
 &= \int_{-\infty}^{+\infty} x \cdot g(x) \cdot dx \int_{-\infty}^{+\infty} y \cdot h(y) \, dy \\
 &= E(X) \cdot E(Y).
 \end{aligned}$$

مثال (٢,٣٥)

لنحسب التوقع الرياضي لحاصل ضرب المتغيرين العشوائيين المستمرين الواردين في المثال (٢,١٦).

الحل

لاحظنا في المثال المذكور أن المتغيرين  $X, Y$  يمثلان متغيرين عشوائيين مستقلين ، وأن دالة كل منهما هي من الشكل :

$$h(y) = \begin{cases} \frac{y^3}{4} & : 0 < y < 2 \\ 0 & : \text{فيما عدا ذلك} \end{cases}$$

$$g(x) = \begin{cases} \frac{1}{2} (1 + 5x^4) & : 0 < x < 1 \\ 0 & : \text{فيما عدا ذلك} \end{cases}$$

كذلك حسبنا في المثال (٢,٣٣) التوقع الرياضي للمتغير  $X$  ، ووجدنا أن :

$$E(X) = \frac{2}{3}$$

لنحسب الآن توقع المتغير  $Y$

$$E(Y) = \int_0^2 y \cdot \frac{y^3}{4} dy = \frac{1}{4} \int_0^2 y^4 dy$$

$$= \frac{8}{5}$$

وبالتعويض في النظرية (٢,٥) نجد أن :

$$E(X) \cdot E(Y) = \frac{2}{3} \cdot \frac{8}{5} = \frac{16}{15}$$

لنحسب الآن دون استخدام فكرة استقلال التوقع الرياضي لجاء الضرب .

$$E(X \cdot Y) = \int_0^1 \int_0^2 x \cdot y \cdot \frac{1}{8} (1 + 5x^4) y^3 dy dx$$

$$= \int_0^1 \frac{x}{8} (1 + 5x^4) dx \cdot \int_0^2 y^4 dy$$

$$= \frac{4}{5} \int_0^1 (x + 5x^5) dx = \frac{4}{5} \left( \frac{1}{2} + \frac{5}{6} \right) = \frac{16}{15}$$

وهي نفس النتيجة التي وجدناها سابقاً .

### (٢,٨) التوقعات الرياضية الخاصة ( التباين — التغير )

#### Special mathematical expectation

تفقدنا النظرية (٢,١) إلى قيمة متوقعة تدعى بالعزم من المرتبة  $K$  حول المبدأ للمتغير العشوائي  $X$  فيما لو عوضنا  $g(x) = x^k$  . نرمز لهذا العزم بالرمز  $\mu_k$  لذلك فإن .

$$\mu_k^1 = E|(X)^k| = \sum x^k \cdot f(x) \quad \text{إذا كان المتغير منقطعاً :}$$

$$= \int x^k \cdot f(x) dx \quad \text{إذا كان المتغير مستمراً :}$$

نلاحظ أنه إذا كان  $K = 0$  فإن  $\mu_0^1 = E(X^0) = 1$  ،

ومنه سواء كان  $X$  مستمراً أو منقطعاً فإن  $E(1) = 1$  .



أما إذا كان  $K = 1$  فإننا نجد أن  $\mu_1' = E(X)$  . سنرمز لهذا العزم بالرمز  $\mu_1$  ، أو بالرمز  $\mu$  . ولذلك فإن :

$$\mu = \mu_1' = E(X)$$

أما إذا عوضنا في النظرية (٢,١) عن  $g(x) = (X - \mu)^k$  فعندئذ تأخذ هذه النظرية الشكل التالي :

$$E(X - \mu)^k = \sum_x (x - \mu)^k f(x) \quad \text{إذا كان } X \text{ منقطعاً} :$$

$$= \int_{-\infty}^{+\infty} (x - \mu)^k \cdot f(x) dx \quad \text{إذا كان } X \text{ مستمراً} :$$

نرمز لهذا العزم بالحديد بالرمز  $\mu_k$  ، ويدعى بالعزم من المرتبة  $K$  حول المتوسط . إن العزم حول المتوسط من المرتبة الثانية والذي نرمز له بالرمز  $\mu_2$  يعطينا معلومات عن تغير العيار حول متوسطه . سنسمى هذا العزم بتباين المتغير العشوائي  $X$  وسنرمز له بالرمز  $\sigma^2$  . وهكذا نجد أن :

$$\sigma^2 = \mu_2 = E(X - \mu)^2$$

يسمى الجذر التربيعي الموجب للتباين بالانحراف المعياري للمتغير العشوائي .

### نظرية (٢,٦)

إن تباين أى متغير عشوائي  $X$  يعطى بالعلاقة :

$$\sigma^2 = E(X^2) - \mu^2$$

### البرهان

من تعريف العزم من المرتبة الثانية حول المتوسط نجد أن :

$$\sigma^2 = E[(x - \mu)^2]$$

$$= E(X^2 + \mu^2 - 2\mu X)$$

$$= E(X^2) + E(\mu^2) - 2\mu E(X)$$

ومن المعلوم أن :

$$E(\mu^2) = \mu^2, \quad E(X) = \mu$$

لذلك فإن :

$$\sigma^2 = E(X^2) - \mu^2$$

مثال (٢,٣٦)

بفرض  $X$  متغير منقطع مُعرّف بالجدول :

$X$	4	10
$f(x)$	$\frac{1}{2}$	$\frac{1}{2}$

وجدنا في المثال (٢,٢٧) أن :

$$E(x) = 7, \quad E(X^2) = 58$$

وحسب النظرية (٢,٦) فإن تباين هذا المتغير يساوى :

$$\sigma^2 = E(X^2) - (E(x))^2$$

$$\sigma^2 = 58 - 49$$

$$\sigma^2 = 9$$

وكذلك فإن الانحراف المعياري لهذا المتغير هو :

$$\sigma = 3$$

مثال (٢,٣٧)

وجدنا في المثال (٢,٢٣) أن للمتغير العشوائى المستمر  $X$  بالكثافة :

$$g(x) = \begin{cases} \frac{1}{2} (1 + 5x^4) & : 0 < x < 1 \\ 0 & : \text{فيما عدا ذلك} \end{cases}$$

توقعاً رياضياً  $E(X) = \frac{2}{3}$  وعزماً من المرتبة الثانية حول المبدأ يساوى  $E(X^2) = \frac{11}{12}$  وحسب النظرية (٢,٦) فإن تباین هذا المتغير يساوى :

$$\sigma^2 = E(X^2) - (E(X))^2 = \frac{11}{12} - \frac{4}{9} = \frac{5}{63}$$

أما انحرافه المعياري المرى إلى جذر تباینه فهو :

$$\sigma = \frac{\sqrt{5}}{\sqrt{63}}$$

إذا عوضنا في التعريف (٢,١٣) عن الدالة :

$$g(x, y) = (x - \mu_x) \cdot (y - \mu_y)$$

حيث يمثل :

$$\mu_y = E(Y), \mu_x = E(X)$$

فإن هذا التعريف يقودنا عندئذ إلى قيمة متوقعة جديدة تدعى بتباین التباين (Covariance) للمتغيرين  $X, Y$  والذي نرمز له بأحد الرمزین  $Cov(X, Y)$  أو  $\sigma_{xy}$ ، لذلك فإن :

$$\sigma_{xy} = E[(X - \mu_x) \cdot (Y - \mu_y)]$$

$$= \sum_x \sum_y (x - \mu_x) \cdot (y - \mu_y) \cdot f(x, y) \quad (١) \quad \text{إذا كان } X, Y \text{ متقطعين}$$

$$(٢) \quad \text{إذا كان } X, Y \text{ مستمرين}$$

$$= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - \mu_x) \cdot (y - \mu_y) f(x, y) dx dy$$

نظرية (٢,٧)

إن تمام تباین متغيرين عشوائيين  $X, Y$  بالمتوسطين  $\mu_y, \mu_x$  على الترتيب يعطى بالعلاقة التالية :

$$\sigma_{xy} = E(X.Y) - \mu_x \cdot \mu_y$$

نلاحظ أن :

$$\begin{aligned}\sigma_{xy} &= E[(X - \mu_x)(Y - \mu_y)] \\ &= E[X.Y - X.\mu_y - Y.\mu_x + \mu_x \cdot \mu_y]\end{aligned}$$

باستخدام النظرية (٢,٢) والنتيجة (٢,٣) نجد أن :

$$\sigma_{xy} = E(X.Y) - E(X.\mu_y) - E(Y.\mu_x) + \mu_x \cdot \mu_y$$

وحيث إن :

$$\mu_y = E(Y), \mu_x = E(X)$$

فإننا نجد أن :

$$\begin{aligned}\sigma_{xy} &= E(X.Y) - \mu_y \cdot \mu_x - \mu_x \cdot \mu_y + \mu_x \cdot \mu_y \\ &= E(X.Y) - \mu_x \cdot \mu_y\end{aligned}$$

مثال (٢,٣٨)

$X, Y$  متغيران عشوائيان مستمران بالكثافة المشتركة :

$$f(x,y) = \begin{cases} 4xy & : 0 < x < 1 \\ & 0 < y < 1 \\ 0 & : \text{فيما عدا ذلك} \end{cases}$$

لنبحث عن تمام المتغيرين  $X, Y$  ، نلاحظ أن :

$$\mu_x = E(X) = \int_{x=0}^1 x \left( \int_{y=0}^1 4xy \, dy \right) dx = \frac{2}{3}$$

$$\mu_y = E(Y) = \int_{y=0}^1 y \left( \int_{x=0}^1 4xy \, dx \right) dy = \frac{2}{3}$$

كما نلاحظ أيضاً أن :

$$E(X, Y) = \int_{y=0}^1 \int_{x=0}^1 4x^2 y^2 dx dy = \frac{4}{9}$$

وأخيراً فإن :

$$\begin{aligned}\sigma_{xy} &= E(X, Y) - \mu_x \cdot \mu_y \\ &= \frac{4}{9} - \left(\frac{2}{3}\right) \cdot \left(\frac{2}{3}\right) \\ &= 0\end{aligned}$$

### (٢,٩) خواص التباين Properties of the variance

فيما يلي سنبرهن أربع نظريات تعتبر مفيدة في حساب التباين أو الانحراف المعياري . سنرمز ب  $\sigma_{g(x)}^2$  و  $\mu_{g(x)}$  للتباين ، ومتوسط الدالة  $g(x)$  على التالى .

#### نظرية (٢,٨)

بفرض  $X$  متغير عشوائى بالتوزيع الاحتمالى  $f(x)$  . عندئذ يكون تباين الدالة  $g(x)$  يعطى بالعلاقة :

$$\sigma_{g(x)}^2 = [E\{g(x) - \mu_{g(x)}\}^2]$$

نلاحظ أن برهان هذه النظرية ينتج مباشرة من تعريف تباين متغير عشوائى وعلى اعتبار أن  $g(x)$  هى دالة بالمتغير العشوائى  $X$  فهو من جديد متغير عشوائى .

#### نظرية (٢,٩)

بفرض  $X$  متغير عشوائى و  $b$  ثابتاً عندئذ يكون :

$$\sigma_{X+b}^2 = \sigma_X^2 = \sigma^2$$

نلاحظ من النظرية السابقة ، وبفرض  $g(X) = X + b$  أن :

$$\sigma_{x+b}^2 = E \{ [(X+b) - \mu_{x+b}]^2 \}$$

ونعلم أن :

$$\mu_{x+b} = E(X + b) = \mu + b$$

$$\sigma_{x+b}^2 = E \{ [(X+b) - (\mu + b)]^2 \} \quad \text{إذا :}$$

$$= E (X - \mu)^2 = \sigma^2$$

وتفسر هذه النظرية أن تبين أى متغير عشوائى لا يتأثر بإضافة أو طرح عدد ثابت إلى هذا المتغير .

### نظرية (٢,١٠)

إذا كان  $X$  متغيرا عشوائيا و  $a$  ثابتا ما ، فعندئذ يكون :

$$\sigma_{aX}^2 = a^2 \sigma_X^2 = a^2 \cdot \sigma^2$$

### البرهان

حسب تعريف التباين نجد أن :

$$\sigma_{a(X)}^2 = E \{ (aX - \mu_{g(x)})^2 \}$$

وحيث إن :

$$\mu_{a(X)} = E(a \cdot X) = a \cdot E(X) = a \cdot \mu$$

لذلك فإن :

$$\sigma_{aX}^2 = E \{ [a(X - \mu)]^2 \}$$

$$= a^2 E (X - \mu)^2$$

$$= a^2 \sigma^2$$

وهذا يعنى أنه إذا ضرب متغير بثابت أو قسم على ثابت فإن التباين يجب أن يضرب بمربع الثابت أو يقسم على مربع الثابت على الترتيب .

## نظرية (٢,١١)

بفرض أن  $X, Y$  متغيرات عشوائيات بالكثافة الاحتمالية المشتركة  $g(x,y)$  عندئذ يكون :

$$\sigma_{aX+bY}^2 = a^2 \sigma_X^2 + b^2 \sigma_Y^2 + 2 a.b.\sigma_{X,Y}$$

## البرهان

من تعريف التباين نجد أن :

$$\sigma_{aX+bY} = E [ \{ (aX + bY) - E(aX + bY) \}^2 ]$$

غير أن :

$$E(aX + bY) = a\mu_X + b\mu_Y$$

$$\sigma_{aX+bY}^2 = E [ \{ (aX + bY) - (a\mu_X + b\mu_Y) \}^2 ] \quad \text{ومنه :}$$

$$= E [ \{ a(X - \mu_X) + b(Y - \mu_Y) \}^2 ]$$

$$= a^2 E(X - \mu_X)^2 + b^2 E(Y - \mu_Y)^2 + 2a.b E(X - \mu_X) \cdot (Y - \mu_Y)$$

$$= a^2 \sigma_X^2 + b^2 \sigma_Y^2 + 2a.b \sigma_{XY}$$

## نتيجة (٢,٤)

إذا كان المتغيران  $X, Y$  مستقلين فعندئذ يكون :

$$\sigma_{aX+bY}^2 = a^2 \cdot \sigma_X^2 + b^2 \cdot \sigma_Y^2$$

بالحقيقة من تعريف تمام التباين  $\sigma_{XY}$  ، والنظرية (٢,٧) نجد ان :

$$\sigma_{XY} = E(X \cdot Y) - \mu_X \mu_Y$$

ومن النظرية (٢,٥) ، وحيث إن المتغيرين مستقلان لذا فإن :

$$\sigma_{XY} = E(X) \cdot E(Y) - \mu_X \cdot \mu_Y = 0$$

وبحسب النظرية (٢,١١) نجد أن :

$$\sigma_{aX+bY}^2 = a^2 \cdot \sigma_X^2 + b^2 \sigma_Y^2 + 2a.b.0$$

وأخيرا فإننا نجد :

$$\sigma_{aX}^2 + bY = a^2 \sigma_X^2 + b^2 \sigma_Y^2$$

نتيجة (٢,٥)

إذا كان  $X, Y$  مستقلين فعندئذ :

$$\sigma_{aX}^2 + bY = a^2 \cdot \sigma_X^2 + b^2 \sigma_Y^2$$

هذه النتيجة تنتج مباشرة بوضع  $b$  عوضا عن  $b$  في النتيجة (٢,٤) .

مثال (٢,٣٩)

بفرض أن تباين المتغيرين  $X, Y$  هما على الترتيب  $\sigma_X^2 = 2$  ،  $\sigma_Y^2 = 4$  ، وبفرض أن تمام تباينها هو  $\sigma_{XY} = -2$  ، فأوجد تباين المتغير  $Z = 5X + 3Y - 4$  .

الحل :

بحسب النظريتين (٢,٩) ، (٢,١١)

$$\sigma_{5x+3y-4}^2 = \sigma_{5x+3y}^2$$

$$\sigma_Z^2 + 25 \sigma_X^2 + 9 \sigma_Y^2 + 2.3.5 \sigma_{X,Y}$$

وبحسب الفرض نجد :

$$\sigma_Z^2 = (25) (2) + (9) (4) + 2.3.5. (-2)$$

$$= 50 + 36 - 60$$

$$= 26$$

(٢,١٠) نظرية تشيبيشيف Chebyshev's theorem

إن الدور الهام الذي يلعبه الانحراف المعياري للمتغير عشوائى كمقياس لتباين قيم المتغير العشوائى عن وسطه ( قيمة المتوقعة ) توضحه لنا متباينة تشيبيشيف الآتية :



## نظرية تشييف

إن احتمال أن يقع متغير عشوائى  $X$  في المجال  $(\mu - k\sigma, \mu + k\sigma)$  أكبر أو يساوى  $(1 - \frac{1}{k^2})$  ، حيث يمثل  $k$  ثابتاً ما ،  $\sigma$  الانحراف المعياري للمتغير  $X$  ،  $\mu$  وسطه .

$$P[\mu - K\sigma < X < \mu + K\sigma] \geq 1 - \frac{1}{K^2}$$

## البرهان

من تعريف تباين متغير عشوائى  $X$  نجد أن :

$$\begin{aligned}\sigma^2 &= E[(X - \mu)^2] = \int_{-\infty}^{+\infty} (x - \mu)^2 \cdot f(x) dx \\ &= \int_{-\infty}^{\mu - k\sigma} (x - \mu)^2 \cdot f(x) dx + \int_{\mu - k\sigma}^{\mu + k\sigma} (x - \mu)^2 \cdot f(x) dx + \int_{\mu + k\sigma}^{+\infty} (x - \mu)^2 f(x) dx.\end{aligned}$$

ومنه :

$$\sigma^2 \geq \int_{-\infty}^{\mu - k\sigma} (x - \mu)^2 f(x) dx + \int_{\mu + k\sigma}^{+\infty} (x - \mu)^2 f(x) dx$$

على اعتبار أن التكامل الأوسط ليس سالباً وبما أن  $|x - \mu| > K\sigma$  في الحالتين أى أن  $(x - \mu)^2 \geq k^2 \sigma^2$  ،  $X \leq \mu - k\sigma$  ،  $X \geq \mu + k\sigma$  فهذا يعنى أن :

$$\sigma^2 \geq \int_{-\infty}^{\mu - k\sigma} k^2 \sigma^2 f(x) dx + \int_{\mu + k\sigma}^{+\infty} k^2 \cdot \sigma^2 f(x) dx$$

ومنه :

$$\frac{1}{k^2} \geq \int_{-\infty}^{\mu - k\sigma} f(x) dx + \int_{\mu + k\sigma}^{+\infty} f(x) dx$$

وأخيراً فإننا نستنتج أن :

$$P[\mu - k\sigma < X < \mu + K\sigma] = \int_{\mu - k\sigma}^{\mu + k\sigma} f(x) dx \geq 1 - \frac{1}{k^2}$$

نلاحظ أنه في الحالة التي يكون فيها  $K = 2$  فإن المتغير  $X$  احتلالاً على الأقل  $1 - \frac{1}{4} = \frac{3}{4}$  للوقوع في المجال  $(\mu - 2\sigma, \mu + 2\sigma)$  . وهذا يعنى أن ثلاث أرباع الملاحظات من أى توزيع ستقع في المجال  $\mu \pm 2\sigma$  . وبشكل مشابه نجد أن ثمانية أتساع الملاحظات  $\frac{8}{9}$  أو أكثر من أى توزيع ستقع في المجال  $\mu \pm 3\sigma$  .

مثال (٢,٤٠)

متغير عشوائي  $X$  توقعه  $\mu = 8$  وتباينه  $\sigma^2 = 9$  ، وتوزيعه غير معروف . لنبحث عن

1.  $P[-4 < X < 20]$
2.  $P[|x - 8| \geq 6]$

الحل

نلاحظ أن :

$$P[-4 < X < 20] = P[8 - 4.3 < X < 8 + 4.3]$$

$$\geq 1 - \frac{1}{4^2} = \frac{15}{16}$$

$$\geq \frac{15}{16}$$

كذلك فإن :

$$\begin{aligned} P[|X - 8| \geq 6] &= 1 - P[|x - 8| < 6] \\ &= 1 - P[8 - 6 < X < 8 + 6] \\ &= 1 - P[8 - 2.3 < X < 8 + 2.3] \\ &\leq 1 - \left(1 - \frac{1}{2^2}\right) \\ &\leq \frac{1}{4} \end{aligned}$$

ملاحظة :

إن النتائج التي تقدمها متباينة تشيبيشيف تكون عادة ضعيفة لأننا نعلم أن احتمال أن يقع المتغير العشوائي  $X$  في المجال  $(\mu - 2\sigma, \mu + 2\sigma)$  ليس أقل من  $\frac{3}{4}$  ، ولكن لا نعلم على وجه الدقة إلى أي حد يمكن أن يكون هذا الاحتمال أكبر من  $\frac{3}{4}$  . أما عندما يكون التوزيع الاحتمالي للمتغير  $X$  معلوماً بالنسبة لنا فنستدثد يمكن أن نحدد بالضبط الاحتمال السابق .

## تمارين محلولة

## تمرين (١)

وضعت نقطة من اللبان على أحد أوجه قطعة نقود بحيث أصبح  $P(H) = \frac{1}{4}$  و  $P(T) = \frac{3}{4}$  ، ثم ألقيت هذه القطعة ثلاث مرات . أوجد جدول قانون توزيع المتغير  $X$  الممثل لعدد الصور التي ظهرت .

## الحل

من الواضح أن فضاء العينة في هذه التجربة هو الفضاء

$$S = \{ (HHH), (HTT), (HHT), (TTH), (HTH), (TTT), (THH), (THT) \}$$

ونلاحظ أن قيم المتغير  $X$  المعرف على الفضاء هي 0, 1, 2, 3 كما نلاحظ أيضا أن :

$$P(HHH) = \frac{1}{4} \cdot \frac{1}{4} \cdot \frac{1}{4} = \frac{1}{64}$$

$$P(HTT) = \frac{1}{4} \cdot \frac{3}{4} \cdot \frac{3}{4} = \frac{9}{64}$$

$$P(HHT) = \frac{1}{4} \cdot \frac{1}{4} \cdot \frac{3}{4} = \frac{3}{64}, P(THT) = \frac{3}{4} \cdot \frac{1}{4} \cdot \frac{3}{4} = \frac{9}{64}$$

$$P(TTH) = \frac{9}{64}, P(TTT) = \frac{3}{4} \cdot \frac{3}{4} \cdot \frac{3}{4} = \frac{27}{64}$$

$$P(HTH) = \frac{3}{64}, P(THH) = \frac{3}{4} \cdot \frac{1}{4} \cdot \frac{1}{4} = \frac{3}{64}$$

ثم أن :

$$f(0) = P(TTT) = \frac{27}{64}$$

$$f(1) = P(HTT) + P(TTH) + P(THT) = \frac{27}{64}$$

$$f(2) = P(HHT) + P(THH) + P(HTH) = \frac{9}{64}$$

$$f(3) = P(HHH) = \frac{1}{64}$$

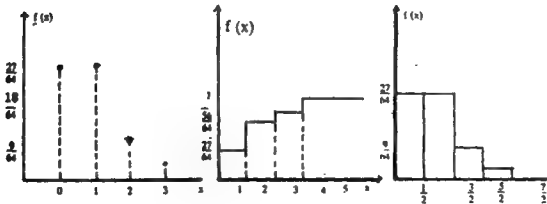
وممكننا نجد أن :

X	0	1	2	3	
f(x)	$\frac{27}{64}$	$\frac{27}{64}$	$\frac{9}{64}$	$\frac{1}{64}$	$\sum_x f(x) = 1$

والملاحظ أن :

$$F(x) = \begin{cases} 0 & \text{من أجل } x < 0 \\ \frac{27}{64} & \text{من أجل } 0 \leq x < 1 \\ \frac{56}{64} & \text{من أجل } 1 \leq x < 2 \\ \frac{63}{64} & \text{من أجل } 2 \leq x < 3 \\ 1 & \text{من أجل } x > 3 \end{cases}$$

أما بالنسبة للمنحنيات الممثلة لدالة التوزيع الاحتمال والتوزيع التراكمي والمضلع التكرارى فهي :



## تمارين (٢)

بفرض أن  $X$  يمثل ضعف الرقم الذي ظهر لدى إلقاء حجر نرد . ما هو جدول قانون توزيع المتغير  $X$  ، ارسم المنحنيات البيانية الموافقة ؟

## الحل

نلاحظ أن عناصر فضاء العينة لتجربة إلقاء حجر النرد مرة واحدة هي :

$$S = \{ 1, 2, 3, 4, 5, 6 \}$$

إذن فقيم المتغير  $X$  هي :

$$X : 2, 4, 6, 8, 10, 12,$$

أما الاحتمالات الموافقة فهي :

$$P(X = 2) = \frac{1}{6}, P(X = 4) = \frac{1}{6}, \dots, P(X = 12) = \frac{1}{6}$$

وجداول قانون التوزيع هو :

x	2	4	6	8	10	12	
f(x)	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\sum f(x) = 1$

كما نلاحظ أن :

$$F(2) = \frac{1}{6}$$

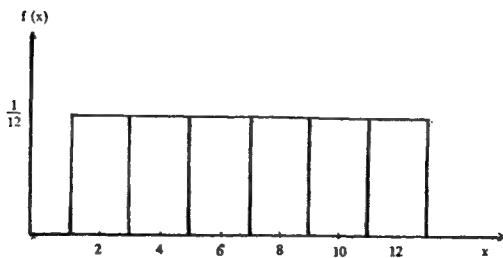
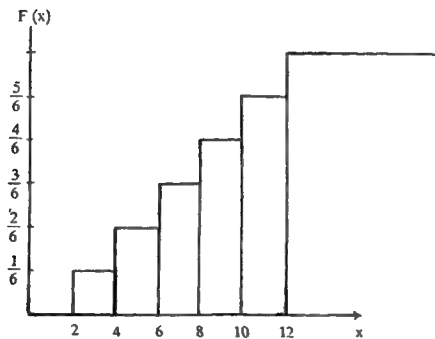
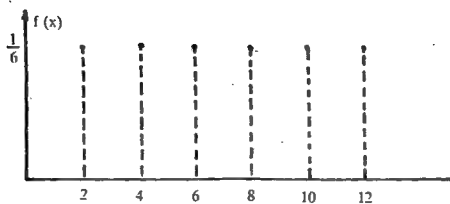
$$F(4) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$$

$$F(6) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}$$

$$F(8) = \frac{4}{6} = \frac{2}{3}$$

$$F(10) = \frac{5}{6}$$

$$F(12) = 1$$



### تمرين (٣)

في المثال (٢، ٨) أوجد  $F(x)$  ، ثم أحسب بواسطته  $P(2.5 \leq x < 4)$  .

#### الحل

من المعلوم أن للمتغير المستمر كثافة احتمالية فوق المجال  $(2, 5)$  وفيما عدا ذلك فإن الكثافة معلومة ، وعلى هذا الأساس . نلاحظ أنه إذا كان  $x \leq 2$  فإن  $F(x) = 0$  أما إذا كان  $2 < x < 5$  فنعدئذ :

$$\begin{aligned} f(x) &= \int_{-\infty}^x f(u) du = \int_{-\infty}^2 f(u) du + \int_2^x f(u) du \\ &= \int_{-\infty}^2 0 \cdot du + \int_2^x \frac{2(1+u)}{27} du = \frac{1}{27} (x^2 + 2x - 8) \end{aligned}$$

أما إذا كان  $x > 5$  فإن :

ولذلك فإن :

$$F(x) = \begin{cases} 0 & \text{if } x \leq 2 \\ 1 & \text{if } x \geq 5 \\ \frac{1}{27} (x^2 + 2x - 8) & \text{if } 2 < x < 5 \end{cases}$$

### تمرين (٤)

بفرض أن  $X$  يمثل مجموع الوجهين اللذين ظهرا لدى إلقاء حجرى نرد متماثلين ومتوازنين ، فتش عن التوزيع الاحتمالى للمتغير  $X$  .

#### الحل

الملاحظ أن قيم المتغير العشوائى المنقطع  $X$  ، والممثل لمجموع الوجهين اللذين ظهرا .

هى :

$$X: 2, 3, 4, 5, \dots, 10, 11, 12$$

أما الاحتمالات الموافقة لهذه القيم فتحدها بواسطة العلاقة :  $f(x) = P[X = x]$  . فمثلا من أجل  $x = 2$  نجد أن :

$$f(2) = P[X = 2] = P[2 = \text{مجموع الرجهين}] = P[1, 1] = \frac{1}{36}$$

كذلك فإن  $f(3)$  يحسب بنفس الطريقة :

$$f(3) = P[X = 3] = P[3 = \text{مجموع الرجهين}] = P[(1,2), (2,1)] = \frac{2}{36}$$

وبنفس الطريق نجد بقية قيم  $f(x)$  . وبترتيب هذه القيم والاحتمالات في جدول نجد أن :

x	2	3	4	5	6	7	8	9	10	11	12
f(x)	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

تحقق من أن  $\sum_x f(x) = 1$  ثم مثل التوزيع السابق بيانيا .

### تمارين (٥)

اشترى سعيد بطاقتين من أصل خمسة آلاف بطاقة من يانصيب خيرى بريال واحد للبطاقة الواحد . ما هو توزيع ربح سعيد في هذه الحالة ؟

### الحل

إذا رمزنا لربح سعيد بالرمز  $X$  ، واعتبرنا أن خسارته هي ربح قدره  $(-2)$  ريالاً فإننا نجد أن للمتغير  $X$  قيمتين فقط هما  $2,4998$  - والاحتمالات الموافقة لهذه القيم هي :

$$f(-2) = P[X = -2] = \frac{4998}{5000} = 0.9996$$

$$f(4998) = P[X = 4998] = \frac{2}{5000} = 0.0004$$

وهكذا نجد أن جدول قانون توزيع ربح سعيد في هذه العملية هو :

X	-2	2998
f(x)	0.9996	0.0004

لاحظ أن  $\sum_x f(x) = 1$



## تمارين (٦)

بفرض  $X$  متغير عشوائي مستمر دالة كثافته من الشكل :

$$f(x) = \begin{cases} \frac{1}{2}x & : 0 \leq x \leq 2 \\ 0 & : \text{فيما عدا ذلك} \end{cases}$$

فالمطلوب تحديد دالة التوزيع  $F(x)$  لهذا المتغير ، ثم رسم المنحنين البيانيين لكل من  $f(x)$  و  $F(x)$  .

## الحل

نلاحظ أن :

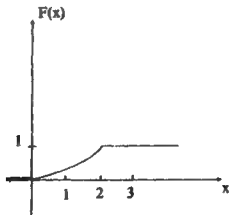
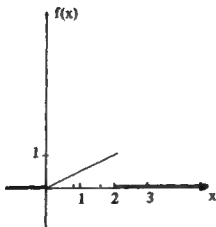
$$F(x) = \int_{-\infty}^x f(u) du$$

وهنا نميز الحالات الثلاث التالية  $0 \leq x \leq 2$  وأخيرا  $x > 2$

ومنه :

$$F(x) = \begin{cases} 0 & : x < 0 \\ \frac{1}{4}x^2 & : 0 \leq x \leq 2 \\ 1 & : x > 2 \end{cases}$$

كما نجد أن



## تمرين (٧)

ألقيت قطعة نقود متماثلة ومتوازنة ثلاث مرات متتالية، فيفرض أن  $X$  يمثل عدد الصور التي ظهرت في الرمية الأولى،  $Y$  عدد الصور التي ظهرت في الرميات الثلاث. أوجد توزيع  $Y, X$ ، ثم التوزيع المشترك للمتغيرين  $(X, Y)$ .

## الحل

نلاحظ أن فضاء العينة يتألف من ثمان نقاط هي :

$$S = (HHH, HTT, THT, TTH, HHT, HTH, THH, TTT)$$

أما قيم المتغير  $X$  فهي 0 أو 1 حسبما يظهر كتابة أو صورة على الترتيب، والاحتمالات الموافقة لهذه القيم فهي  $\frac{1}{2}$ ،  $\frac{1}{2}$  على التتالي نظراً لتوازن وتمائل قطعة النقود. وهكذا نجد أن جدول توزيع المتغير  $X$  هو :

X	0	1
g(x)	$\frac{1}{2}$	$\frac{1}{2}$

كما نلاحظ أن قيم  $Y$  الممثل لعدد الصور التي ظهرت في الرميات الثلاث هي :

$$Y = \{0, 1, 2, 3\} \dots$$

أما الاحتمالات الموافقة فتحسب بالعلاقة :

$$h(y) = P [ Y = y ]$$

والملاحظ أن :

$$h(0) = P [ (TTT) ] = \frac{1}{8}$$

$$h(1) = P [ (HTT, THT, TTH) ] = \frac{3}{8}$$

$$h(2) = \frac{3}{8}, h(3) = \frac{1}{8}$$

وبترتيب هذه القيم بمجدول نجد أن :

Y	0	1	2	3
h(y)	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$

أما التوزيع الاحتمالي المشترك لـ (X, Y) فهو :

Y \ X	0	1	المجموع
0	$\frac{1}{8}$	0	$\frac{1}{8}$
1	$\frac{2}{8}$	$\frac{1}{8}$	$\frac{3}{8}$
2	$\frac{1}{8}$	$\frac{2}{8}$	$\frac{3}{8}$
3	0	$\frac{1}{8}$	$\frac{1}{8}$
المجموع	$\frac{4}{8}$	$\frac{4}{8}$	1

تمارين (٨)

متغير عشوائي Y دالة كثافته من الشكل :

$$f(y) = \begin{cases} \alpha & : a \leq y \leq b \\ 0 & : \text{فيما عدا ذلك} \end{cases}$$

(١) حدد  $\alpha$

(٢) أوجد دالة التوزيع Y (F(y))

(٣) مثل بيانياً كلاً من :  $F(y), f(y)$

الحل

من خواص دالة الكثافة نعلم أن :  $\int_{-\infty}^{+\infty} f(y) dy = 1$

لذلك فإن

$$\int_0^1 \alpha \cdot dy = 1$$

ومنه :

$$\alpha = \frac{1}{b-a}$$

وبالتعويض في عبارة دالة الكثافة نجد أنها تأخذ الشكل التالي :

$$f(y) = \begin{cases} \frac{1}{b-a} & : a \leq y \leq b \\ 0 & : \text{فيما عدا ذلك} \end{cases}$$

من تعريف دالة التوزيع  $F$  نجد أن :

$$F(y) = \int_{-\infty}^y f(u) du$$

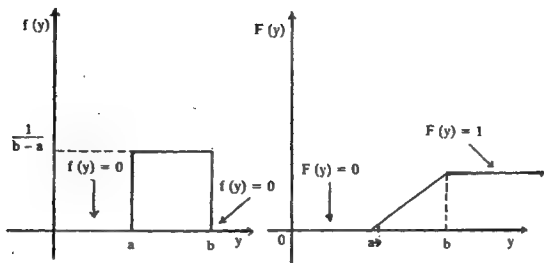
وهنا نميز الحالات الثلاث التالية :

$$y > b, a \leq y \leq b, y < a$$

ف نجد أن :

$$F(y) = \begin{cases} 0 & : y < a \\ \frac{y-a}{b-a} & : a \leq y \leq b \\ 1 & : y > b \end{cases}$$

وأخيراً فإننا نجد أن :



تمارين (٩)

بفرض أن الكثافة المشتركة للمتغيرين العشوائيين  $(X, Y)$  هي من الشكل :

$$f(x, y) = \begin{cases} 8xy & : 0 < x < 1 \\ & : 0 < y < x \\ 0 & : \text{فيما عدا ذلك} \end{cases}$$

فالمطلوب البحث عن :

$$P \left[ Y < \frac{1}{8} \mid X = \frac{1}{2} \right], f(y/x), h(y), g(x)$$

الحل

نلاحظ أن :

$$g(x) = \int_{-\infty}^{+\infty} f(x, y) dy = \int_0^x 8x \cdot y dy$$

ومنه :

$$g(x) = \begin{cases} 4x^3 & : 0 < x < 1 \\ 0 & : \text{فيما عدا ذلك} \end{cases}$$

كذلك فإن :

$$h(y) = \int_{-\infty}^{+\infty} f(x, y) dx = \int_y^1 8xy dx$$

ومنه :

$$h(xy) = \begin{cases} 4y(1-y^2) & : 0 < y < 1 \\ 0 & : \text{فيما عدا ذلك} \end{cases}$$

ولكن نعلم أن :

$$f(y|x) = \frac{f(x, y)}{g(x)}$$

ولذلك فإن :

$$f(y|x) = \frac{2y}{x^2} : 0 < y < x, 0 < x < 1$$

وكذلك فإن :

$$P[Y < 1/2 | X = 1/2] = \int_0^{1/2} 8y dy = \frac{1}{16}$$

تمارين (١٠)

X, Y متغيران عشوائيان بالتوزيع المشترك :

X \ Y	X			المجموع
	2	3	4	
1	0.06	0.15	0.09	0.3
2	0.14	0.35	0.21	0.7
المجموع	0.2	0.5	0.3	1

ابحث فيما إذا كان المتغيران السابقان مستقلين أم لا ؟

الحل :

نلاحظ أن جدول توزيع كل من المتغيرين  $Y, X$  هو من الشكل :

X	2	3	4
g(x)	0.2	0.5	0.3

X	1	2
g(X)	0.3	0.7

كما نلاحظ أن :

$$g(2) \cdot h(1) = 0.2 \times 0.3 = 0.06 = f(2,1)$$

$$g(2) \cdot h(2) = 0.2 \times 0.7 = 0.14 = f(2,2)$$

$$g(3) \cdot h(1) = 0.5 \times 0.3 = 0.15 = f(3,1)$$

$$g(3) \cdot h(2) = 0.5 \times 0.7 = 0.35 = f(3,2)$$

$$g(4) \cdot h(1) = 0.3 \times 0.3 = 0.09 = f(4,1)$$

$$g(4) \cdot h(2) = 0.3 \times 0.7 = 0.21 = f(4,2)$$

من العلاقات السابقة نجد أنه من أجل أى  $(x,y)$  فإن  $f(x,y) = g(x) \cdot h(y)$  مما ينتج معه أن المتغيرين السابقين مستقلان .

تموين (١١)

بفرض أن التوزيع المشترك للمتغيرين  $X, Y$  هو من الشكل :أوجد توزيع كلا من  $Y, X$  والتوزيعات الشرطية ؟

Y \ X	X			المجموع
	1	2	3	
1	0	$\frac{1}{6}$	$\frac{1}{12}$	$\frac{3}{12}$
2	$\frac{1}{5}$	$\frac{1}{9}$	0	$\frac{14}{45}$
3	$\frac{2}{15}$	$\frac{1}{4}$	$\frac{1}{18}$	$\frac{79}{180}$
المجموع	$\frac{5}{15}$	$\frac{19}{36}$	$\frac{5}{36}$	1

الحل

نلاحظ أن توزيع المتغير X هو :

X	1	2	3
g(x)	$\frac{1}{3}$	$\frac{19}{36}$	$\frac{5}{36}$

كما أن توزيع Y هو من الشكل :

Y	1	2	3
h(y)	$\frac{1}{4}$	$\frac{14}{45}$	$\frac{79}{180}$

ثم إن التوزيع الشرطي للمتغير X بمعلومية أن Y قد أخذ قيمة يحسب بالعلاقة :



$$f(X | Y = y) = \frac{f(x, y)}{h(y)}$$

فمثلاً :

$$f(1 | 1) = \frac{f(1, 1)}{h(1)} = \frac{0}{1/4} = 0$$

$$f(2 | 1) = \frac{f(2, 1)}{h(1)} = \frac{1/6}{1/4} = \frac{2}{3}$$

$$f(3 | 1) = \frac{f(3, 1)}{h(1)} = \frac{1/12}{1/4} = \frac{1}{3}$$

$$f(2 | 1) = \frac{f(2, 1)}{h(1)} = \frac{1/6}{1/4} = \frac{2}{3}$$

وهكذا بالنسبة لبقية القيم .

وبنفس الطريقة نحسب التوزيع الشرطي للمتغير  $Y$  بمعلومية أن  $X$  قد أخذ قيمة :

$$f(Y | X = x) = \frac{f(x, y)}{g(x)}$$

فمثلاً نجد أن :

$$f(1 | 3) = \frac{f(3, 1)}{g(3)} = \frac{1/12}{5/36} = \frac{3}{5}$$

$$f(2 | 3) = \frac{f(3, 2)}{g(3)} = \frac{0}{5/36} = 0$$

وهكذا ....

تمرين (١٢)

بضاعة مصنعة مؤلفة من 20 قطعة مخلوطة خلطاً جيداً بينها قطعتين معابتين .  
اخترنا منها أربع قطع بصورة عشوائية ماهى القيمة المتوقعة  $\mu$  لعدد القطع المعابة ؟

الحل

لنرمز لعدد القطع المعابة والموجودة في العينة المسحوبة بالرمز  $X$  . نلاحظ أن قيم هذا المتغير هي 0, 1, 2 . كما نلاحظ أن فضاء العينة  $S$  في عملية سحب أربع قطع من بين 20 قطعة يتكون من :

$\binom{20}{4} = 4845$  قطعة ، كما نلاحظ أيضا أن الاحتمال الموافق لعدد القطع المعابة المسحوبة  $X$  هو :

$$P[X = x] = \frac{\binom{2}{x} \binom{18}{4-x}}{\binom{20}{4}}$$

وهكذا نجد أن القيمة المتوقعة لعدد القطع المعابة هو :

$$\begin{aligned} \mu = E(X) &= \sum_{x=0}^2 x \cdot P[X = x] \\ &= 0 \cdot \frac{\binom{2}{0} \binom{18}{4}}{\binom{20}{4}} + 1 \cdot \frac{\binom{2}{1} \binom{18}{3}}{\binom{20}{4}} + 2 \cdot \frac{\binom{2}{2} \binom{18}{2}}{\binom{20}{4}} \end{aligned}$$

ومنه :

$$\mu = 0 + 1 \cdot \frac{2(816)}{4845} + 2 \cdot \frac{1(153)}{4845}$$

$$\mu = 1938/4845 = 0.4$$

تمرين (١٣)

ابحث عن التوقع الرياضي لكل من المتغيرات  $X, Y, Y, X$  الواردة في التمرين (٧، ١٠) ثم أثبت أن توقع  $X, Y$  يساوى حاصل جداء توقع  $X$  في توقع  $Y$

الحل

من جدول توزيع  $X$  نجد أن :

$$\mu_x = \sum_x x \cdot g(x) = 2(0.2) + 3(0.5) + 4(0.3) = 3.1$$

ومن جدول توزيع المتغير  $Y$  نجد أيضا أن :

$$\mu_y = \sum_y y \cdot h(y) = 1(0.3) + 2(0.7) = 1.7$$

ومن المعلوم أن :

$$\begin{aligned}\mu_{x,y} &= E[X, Y] \\ &= \sum_x \sum_y x \cdot y \cdot f(x, y)\end{aligned}$$

ومن جدول التوزيع  $Y, X$  المشترك نجد كذلك أن :

$$\begin{aligned}\mu_{x,y} &= (2.1) f(2,1) + (2.2) f(2,2) + \\ &\quad (3.1) f(3,1) + (3.2) f(3,2) + \\ &\quad (4.1) f(4,1) + (4.2) f(4,2)\end{aligned}$$

وبالتعويض عن  $f(x,y)$  بقيمتها من الجدول نجد :

$$\begin{aligned}\mu_{x,y} &= (2.1) (0.06) + (2.2) (0.14) + \\ &\quad (3.1) (0.15) + (3.2) (0.35) + \\ &\quad (4.1) (0.09) + (4.2) (0.21) \\ &= 0.12 + 0.56 + 0.45 + 2.1 + 0.36 + 1.68 \\ \mu_{x,y} &= 5.27\end{aligned}$$

نلاحظ مما سبق أن :

$$\mu_x \cdot \mu_y = (3.1) \cdot (1.7) = 5.27$$

وبمقارنة النتيجةين الأخيرتين نجد المطلوب .

#### تمارين (١٤)

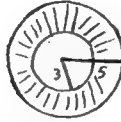
يرمى رامي على هدف مكون من دائرتين متمركزتين نصفى قطريهما على الترتيب 3, 5 فإذا أصاب الرامي الدائرة الداخلية فإنه يسجل له العدد 10 ، أما إذا أصاب القسم المظلل فإنه يسجل له 5 . فإذا ، علمت أن احتمال إصابته للهدف هو 0.70 ، وأن احتمالات إصابته لأى نقطة من الهدف متساوية . فما هى القيمة المتوقعة  $\mu$  لعدد النقاط التى يسجلها فى كل رمية ؟

#### الحل

من الملاحظ أن احتمال حصول الرامي على 10 أو 5 من النقاط على الترتيب هو :

$$f(10) = 0.7 \frac{\text{مساحة الدائرة الصغرى}}{\text{مساحة الهدف}} = 0.7 \frac{3^2}{5^2} = 0.252$$

$$f(5) = 0.7 \frac{\text{مساحة الحلقة المظللة}}{\text{مساحة الهدف}} = 0.7 \frac{(5^2 - 3^2)}{5^2} = 0.448$$



ونلاحظ أن القيمة المتوقعة لعدد النقاط التي يسجلها في كل رمية هي :

$$\mu = 5 \cdot (0.448) + 10 \cdot (0.252) = 4.76$$

تمرين (١٥)

ما هو التوقع الرياضي للعدد الملاحظ على حجر نرد عند إلقائه ؟

الحل

إذا رمزنا للعدد الملاحظ بالرمز  $X$  ، فإننا نجد أن جدول توزيع  $X$  هو :

$X$	1	2	3	4	5	6
$f(x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

ومنه :

$$\begin{aligned} \mu = E(X) &= \frac{1}{6} \cdot 1 + \frac{1}{6} \cdot 2 + \frac{1}{6} \cdot 3 + \frac{1}{6} \cdot 4 + \frac{1}{6} \cdot 5 + \frac{1}{6} \cdot 6 \\ &= 3.5 \end{aligned}$$

(١٦) تمّين

بفرض أن  $X$  يمثل الوجه الذى ظهر عند إلقاء حجر نرد متوازن ومتماثل . أوجد توقع المتغير  $Y = 2X^2 - 5$  ؟

الحل

نعلم أن جدول توزيع المتغير  $X$  هو من الشكل :

$X$	1	2	3	4	5	6
$f(x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

كما أن :

$X^2$	1	4	9	16	25	36
$f(x^2)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

والملاحظ أن جدول توزيع المتغير  $Y$  هو من الشكل :

$Y = 2x^2$	-3	3	13	27	45	67
$f(2x^2)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

وهكذا نجد أن :

$$E[Y] = \frac{1}{6} [-3 + 3 + 13 + 27 + 45 + 67] = \frac{1}{6} (152)$$

وأخيراً فإن :

$$E[Y] = \frac{76}{3}$$

## تمرين (١٧)

في المثال (٢،٢١) احسب التوقع الرياضي للمتغير  $X^2 - 1$ ، ثم احسب تباين المتغير  $X$ .

## الحل

نعلم أن جدول توزيع المتغير  $X$  هو من الشكل :

X	-1	0	1	2
f(x)	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{3}{8}$	$\frac{1}{4}$

نلاحظ أن :

$$E(X^2 - 1) = \sum_x (x^2 - 1) f(x)$$

$$= 0 \cdot \frac{1}{8} + (-1) \cdot \frac{1}{4} + 0 \cdot \frac{3}{8} + 3 \cdot \frac{1}{4} = \frac{1}{2}$$

ومن خواص التوقع الرياضي نجد أن :

$$E(X^2) - 1 = \frac{1}{2}$$

ومنه :

$$E(X^2) = \frac{3}{2}$$

وأخيراً فإن تباين المتغير  $X$  يعطى بالملاقة :

$$\sigma^2 = E(X^2) - \mu^2 = \frac{3}{2} - \left(\frac{3}{4}\right)^2$$

$$\sigma^2 = 1.3125$$

هنا عوضنا عن  $\mu = \frac{3}{4}$  كما وجدنا في المثال (٢،٢١)

## تمرين (١٨)

جهاز إرسال يحوى سبع ترانزستورات ، اثنان منها عاطلان . اخترنا بصورة عشوائية ثلاثة ترانزستورات ، ونزعناها من جهاز الإرسال ، وفحصناها . ما هو التوقع الرياضى لعدد الترانزستورات العاطلة المنزوعة ؟

## الحل

لنرمز بـ  $Y$  لعدد الترانزستورات العاطلة المنزوعة .

نلاحظ أن جدول قانون توزيع  $Y$  هو من الشكل :

X	0	1	2
f(x)	$\frac{10}{35}$	$\frac{20}{35}$	$\frac{5}{35}$

ذلك لأن :

$$P[Y = y] = \frac{\binom{2}{y} \binom{5}{3-y}}{\binom{7}{3}}$$

وهكذا نجد أن توقع  $Y$  هو :

$$\mu_y = E(Y) = 0. \left(\frac{10}{35}\right) + 1. \left(\frac{20}{35}\right) + 2. \left(\frac{5}{35}\right) = 0.857$$

ولحساب التباين نعلم أن :

$$\sigma_y^2 = E(Y^2) - \mu_y^2$$

غير أن :

$$E(Y^2) = \sum_y y^2 \cdot h(y)$$

$$= 0. \left(\frac{10}{35}\right) + 1. \left(\frac{20}{35}\right) + 4. \left(\frac{5}{35}\right)$$

$$= \frac{40}{35} = 1.1428$$

بالتعويض نجد أن :

$$\begin{aligned}\sigma_y^2 &= 1.1428 - (0.857)^2 \\ &= 0.408\end{aligned}$$

تمارين (١٩)

X متغير عشوائى مستمر بالكثافة الاحتمالية :

$$f(x) = \begin{cases} k e^{-x} & : x \geq 0 \\ 0 & : \text{فيما عدا ذلك} \end{cases}$$

والمطلوب :

١ — البحث عن قيمة الثابت K ؟

٢ — توقع المتغير X ؟

٣ — تبين هذا المتغير ؟

الحل

نلاحظ من خواص الكثافة الاحتمالية أن :

$$\int_{-\infty}^{+\infty} f(x) dx = 1 \Leftrightarrow \int_0^{\infty} K e^{-x} dx = 1 = K = 1$$

والكثافة الاحتمالية بعد تعويض K = 1 هي من الشكل :

$$f(x) = \begin{cases} e^{-x} & : x \geq 0 \\ 0 & : \text{فيما عدا ذلك} \end{cases}$$

أما التوقع الرياضى في هذه الحالة فيحسب بالعلاقة :



$$\mu = E(X) = \int_{-\infty}^{+\infty} x \cdot f(x) dx = \int_0^{\infty} x \cdot e^{-x} dx$$

وبالمكاملة بالتجزئة باعتبار أن  $u = x$ ,  $dv = e^{-x} dx$  نجد أن :

$$\mu = -x \cdot e^{-x} \Big|_{x=0}^{\infty} + \int_0^{\infty} e^{-x} \cdot dx$$

$$\mu = 0 + 1 = 1$$

ولحساب التباين نبدأ بحساب  $\mu_2 = E(X^2)$  فنجد أن :

$$\mu_2 = E(X^2) = \int_0^{\infty} x^2 \cdot e^{-x} dx$$

وبالمكاملة بالتجزئة على أساس أن  $u = x^2$ ,  $dv = e^{-x} dx$  نجد أن :

$$\mu_2 = x^2 \cdot e^{-x} \Big|_{x=0}^{\infty} + 2 \int_0^{\infty} x e^{-x} dx$$

ومن الملاحظ أن  $\mu = \int_0^{\infty} x e^{-x} dx = 1$  وهكذا نجد :

$$\mu_2 = E(X^2) = 0 + 2 \cdot 1 = 2$$

وأخيراً فإن :

$$\sigma^2 = \mu_2 - \mu^2 = 2 - 1 = 1$$

تمارين (٢٠)

متغير عشوائى  $Y$  توقعه الرياضى  $\mu = 10$  ، وانحرافه المعيارى  $\sigma = 2$  احسب باستخدام متباينة تشيبيشيف :

$$P[|Y - 10| \geq 3] \quad \text{—} \quad ١$$

$$P[|Y - 10| < 3] \quad \text{—} \quad ٢$$

$$P[5 < Y < 15] \quad \text{—} \quad ٣$$

الحل

أولاً — نلاحظ أن :

$$\begin{aligned}
 P[|Y - 10| \geq 3] &= 1 - P[|X - 10| < 3] \\
 &= 1 - P[10 - 3 < X < 10 + 3] \\
 &= 1 - P\left[10 - \frac{3}{2} \cdot 2 < X < 10 + \frac{3}{2} \cdot 2\right]
 \end{aligned}$$

ولكن نعلم أن :

$$P[\mu - k \cdot \sigma < X < \mu + k \cdot \sigma] \geq 1 - \frac{1}{k^2}$$

وبملاحظة أن  $k = \frac{3}{2}$  فإننا نجد أن :

$$\begin{aligned}
 P[|Y - 10| \geq 3] &\leq 1 - 1 + \frac{1}{k^2} \\
 &\leq \frac{4}{9}
 \end{aligned}$$

ثانياً — نعلم أن :

$$P[|Y - 10| < 3] = 1 - P[|Y - 10| \geq 3]$$

وباستخدام نتيجة الطلب الأول نجد أن :

$$P[|Y - 10| \geq 3] \geq \frac{5}{9}$$

ثالثاً — نلاحظ أن :

$$P[5 < Y < 15] = P[10 - 5 < Y < 10 + 5] \geq 1 - \frac{1}{k^2}$$

حيث يمثل  $K \cdot \sigma = 2K = 5$  ومنه  $K = \frac{5}{2}$

وأخيراً فإن :

$$P[5 < Y < 15] \geq \frac{21}{25}$$

## تمارين (٢١)

احسب  $P [\mu - 2\sigma < X < \mu + 2\sigma]$  إذا علمت أن للمتغير  $X$  توزيعاً معطى بالكثافة :

$$f(x) = \begin{cases} 6x(1-x) & : 0 < x < 1 \\ 0 & : \text{فيما عدا ذلك} \end{cases}$$

## الحل

نبدأ بحساب  $\mu$ ,  $E(X^2)$ , ثم  $\sigma^2$  فنجد أن :

$$\mu = \int_0^1 x \cdot [6x(1-x)] dx = \frac{1}{2}$$

$$E(X^2) = \int_0^1 x^2 [6x(1-x)] dx$$

$$E(X^2) = \frac{3}{10}$$

وهكذا نجد أن :

$$\sigma^2 = E(X^2) - \mu^2 = \frac{3}{10} - \frac{1}{4} = \frac{1}{20}$$

وأخيراً فإن :

$$2\sigma = \frac{1}{\sqrt{5}} = 0.4472$$

ومنه :

$$P [\mu - 2\sigma < X < \mu + 2\sigma] = P [0.5 - 0.447 < X < 0.5 + 0.447]$$

$$P [0.053 < X < 0.947]$$

$$\int_{0.053}^{0.947} 6x(1-x) dx$$

$$= 0.98$$

تمرين (٢٢)

متغيران عشوائيان منقطعان جدول توزيعهما المشترك من الشكل :

$\begin{matrix} X \\ Y \end{matrix}$	-2	-1	0	1	2	3	
0	0.05	0.05	0.1	0	0.05	0.05	0.30
1	0.10	0.05	0.05	0.1	0	0.05	0.35
2	0.03	0.12	0.07	0.06	0.03	0.04	0.35
	0.18	0.22	0.22	0.16	0.08	0.14	1

والمطلوب :

١ — إيجاد توزيع  $Z = X + Y$

٢ — حساب التوقع الرياضي لـ  $Z$  والتحقق من أن  $E(Z) = E(X) + E(Y)$

الحل

من الواضح أنه إذا رمزنا بـ  $\phi$  للتوزيع الاحتمالي للمتغير  $Z$  لوجدنا أن قيم هذا المتغير هي :

$$Z: -2, -1, 0, 1, 2, 3, 4, 5$$

كما أن قيم التوزيع الاحتمالي في هذه النقاط هي :

$$\phi(-2) = 0.05$$

$$\phi(-1) = 0.05 + 0.1 = 0.15$$

$$\phi(0) = 0.1 + 0.05 + 0.03 = 0.18$$

$$\phi(1) = 0 + 0.05 + 0.12 = 0.17$$

$$\phi(2) = 0.05 + 0.1 + 0.07 = 0.22$$

$$\phi(3) = 0.05 + 0 + 0.06 = 0.11$$

$$\phi(4) = 0.05 + 0.03 = 0.08$$

$$\phi(5) = 0.04$$

وهكذا نستنتج أن جدول توزيع المتغير الجديد Z هو من الشكل :

Z	-2	-1	0	1	2	3	4	5
$\phi(Z)$	0.05	0.15	0.18	0.17	0.22	0.11	0.08	0.04

أما التوقع الرياضي للمتغير Z فيحسب بالشكل التالي :

$$\mu_z = E(Z) = (-2)(0.05) + (-1)(0.15) + (0)(0.18) + (1)(0.17) + (2)(0.22) + (3)(0.11) + (4)(0.08) + (5)(0.04)$$

ومنه :

$$\mu_z = 1.21$$

غير أن :

$$\mu_x = (-2)(0.18) + (-1)(0.22) + (0)(0.22) + (1)(0.16) + (2)(0.08) + (3)(0.14)$$

ومنه :

$$\mu_x = 0.16$$

كذلك فإن :

$$\mu_y = (0)(0.3) + (1)(0.35) + (2)(0.35) = 1.05$$

وأخيراً فإننا نلاحظ أن :

$$\mu_z = \mu_x + \mu_y$$

### تمارين (٢٣)

يجرى الرمي على هدف بواسطة سلاح معين . فإذا علمت أن احتمال إصابة الهدف  $P$  ، وأن  $X, Y$  يمثلان على الترتيب عدد مرات أخطاء ، واصابة الهدف ، فالمطلوب إيجاد دالة التوزيع المشتركة لهذين المتغيرين ؟

### الحل

نلاحظ أن المتغيرين العشوائيين  $X, Y$  مرتبطان بالعلاقة التابعة :

$$X + Y = 1$$

لذلك فإن التوزيع المشترك لهذين المتغيرين يوضحه الجدول التالي :

Y \ X	0	1
0	0	q
1	p	0

### تمارين (٢٤)

يصوب راميان على هدف ل كل بسلاحه وبصورة مستقلة عن الآخر ، فإذا علمت أن عدد المرات التي أصاب بها الرامي الأول الهدف هو  $X$  ، والرامي الثاني هو  $Y$  ، وأن احتمال إصابة الهدف من قبل الرامي الأول و  $P_1$  من قبل الرامي الثاني ، فأوجد دالة التوزيع المشترك للمتغيرين  $X, Y$  ؟

### الحل

بما أن كل رام يصوب على الهدف بصورة مستقلة عن الآخر لذلك فإن :

$$\begin{aligned} F(x, y) &= P(X < x) \cdot P(Y < y) \\ &= F_1(x) \cdot F_2(y) \end{aligned}$$

ولكن :

$$F_1(x) = \begin{cases} 0 & : x \leq 0 \\ q_1 = 1 - p_1 & : 0 < x \leq 1 \\ 1 & : x > 1 \end{cases}$$

وكذلك فإن :

$$F_2(y) = \begin{cases} 0 & : y \leq 0 \\ q_2 = 1 - p_2 & : 0 < y \leq 1 \\ 1 & : y > 1 \end{cases}$$

لذلك فإن جدول التوزيع المشترك لهذين المتغيرين هو :

Y \ X	X		
	$x \leq 0$	$0 < x \leq 1$	$x > 1$
$y \leq 0$	0	0	0
$0 < y \leq 1$	0	$q_1 q_2$	$q_2$
$1 < y$	0	$q_1$	1

تمارين (٢٥)

يجرى التصويب على هدف بـ ٢ سلاحين مستقلين واحتمال إصابة الهدف بكل سلاح هو  $p$  ، فإذا علمت أن  $X$  يمثل الفرق بين عدد الإصابات وعدد الطلقات التي أخطأت الهدف ، وأن  $Y$  يمثل مجموع عدد الإصابات وعدد الطلقات التي أخطأت الهدف فالمطلوب :

(١) البحث عن توزيع كل من  $Y, X$

(٢) حساب  $D(Y)$ ,  $D(X)$ ,  $E(Y)$ ,  $E(X)$  ؟

الحل

نلاحظ أن :

X	-2	0	2
P(x)	$q^2$	$2pq$	$p^2$

Y	2
P(Y)	1

لذلك وحسب تعريف التوقع الرياضي ، والتباين نجد أن :

$$E(X) = 2(p^2 - q^2) = 2(p - q)$$

$$E(Y) = 2(1) = 2$$

$$D(X) = 4(p^2 + q^2) - 4(p^2 - 2pq + q^2) = 8pq$$

$$D(Y) = 0$$

تمرين (٢٦)

يرمى راميان على هدف معين كل بسلاحه وبصورة مستقلة عن الآخر . فإذا علمت أن احتمال إصابة الأول للهدف هو  $P_1$  ، والثاني هو  $P_2$  وأن  $X$  يمثل عدد المرات التي أصاب بها الأول الهدف ،  $Y$  عدد المرات التي أصاب بها الثاني الهدف ، وأن  $Z$  يمثل الفرق بين عدد مرات الإصابة . فالمطلوب هو إيجاد قانون توزيع  $Z$  وتوقعه الرياضي

الحل

نلاحظ أن للمتغير  $Z$  ثلاث قيم هي  $1, 0, -1$  ، وأن :



$$P(Z = -1) = P(X = 0) P(Y = 1) = q_1 p_2$$

$$P(Z = 0) = P(X = 0) P(Y = 0) + P(X = 1) P(Y = 1) = q_1 q_2 + p_1 p_2$$

$$P(Z = 1) = P(X = 1) P(Y = 0) = p_1 q_2$$

Z	-1	0	1
P(z)	$q_1 p_2$	$q_1 q_2 + p_1 p_2$	$p_1 q_2$

من الجدول السابق نستنتج أن :

$$E(Z) = -q_1 p_2 + p_1 q_2 = p_1 - p_2$$

تمرين (٢٧)

بفرض أن دالة الكثافة المشتركة للمتغيرين المستمرين  $X, Y$  هي من الشكل :

$$f(x, y) = \begin{cases} \frac{1}{\pi r^2} & : x^2 + y^2 < r \\ 0 & : x^2 + y^2 > r \end{cases}$$

فأوجد دالة الكثافة الشرطية للمتغير  $X$  ، علماً بأن المتغير  $Y$  قد أخذ قيمة  $y$  وأن

$$0 < |y| < r$$

الحل

من المعلوم أن :

$$f(x|Y=y) = \frac{f(x, y)}{f(y)}$$

ومن الفرض لدينا :

$$= \frac{\frac{1}{\pi r^2}}{\frac{1}{\sqrt{r^2 - y^2}}} = \frac{1}{\pi r^2} \int_{-\sqrt{r^2 - y^2}}^{\sqrt{r^2 - y^2}} dx$$

$$= \frac{1}{2\sqrt{r^2 - y^2}} \quad : \quad |x| < \sqrt{r^2 - y^2}$$

وبما أن :

$$f(x, y) = 0 \text{ for } x^2 + y^2 > r^2$$

لذلك فإن :

$$f(x | Y = y) = 0 \quad \text{for} \quad |x| > \sqrt{r^2 - y^2}$$

تمرين (٢٨)

بفرض أن دالة الكثافة المشتركة للثنائية (X, Y) هي من الشكل :

$$f(x, y) = \begin{cases} \frac{1}{6\pi} & : \frac{x^2}{9} + \frac{y^2}{4} < 1 \\ 0 & : \frac{x^2}{9} + \frac{y^2}{4} > 1 \end{cases}$$

فابحث عن دالة الكثافة لكل من Y, X ؟

الحل :

$$f(x) = \int_{-\infty}^{+\infty} f(x, y) dy = \frac{1}{6\pi} \int_{-2\sqrt{1 - \frac{x^2}{9}}}^{2\sqrt{1 - \frac{x^2}{9}}} dy$$

$$g(x) = \begin{cases} \frac{2}{9\pi} \sqrt{9 - x^2} & : |x| < 3 \\ 0 & : |x| \geq 3 \end{cases} \quad \text{وبنه}$$

بنفس الطريقة نجد أن :

$$h(y) = \begin{cases} \frac{1}{2\pi} \sqrt{4 - y^2} & : |y| < 2 \\ 0 & : |y| \geq 2 \end{cases}$$

## تمرين (٢٩)

بفرض أن الجدول الثاني يمثل التوزيع المشترك للمتغيرين  $X, Y$  :

$\begin{matrix} X \\ Y \end{matrix}$	$x_1$	$x_2$	$x_3$
$y_1$	0.1	0.3	0.2
$y_2$	0.6	0.18	0.16

أوجد دالة الكثافة الشرطية للمتغير  $X$  إذا علمت أن المتغير  $Y$  قد أخذ القيمة  $y_1$  .

الحل

نعلم أن :

$$f(x_1 | Y = y_1) = \frac{P[X = x_1, Y = y_1]}{h(y_1)}$$

وبالتعويض نجد أن :

$$f(x_2 | Y = y_1) = \frac{0.1}{0.6} = \frac{1}{6}$$

كذلك فإن :

$$f(x_2 | Y = y_1) = \frac{0.3}{0.6} = \frac{1}{2}$$

$$f(x_3 | Y = y_1) = \frac{0.2}{0.6} = \frac{1}{3}$$

## تمارين عامة

- (١) حدد فيما اذا كانت المتغيرات التالية متقطعة أم مستمرة :
- أ -  $X$  - عدد حوادث السيارات التي وقعت خلال عام في المملكة العربية السعودية .
- ب -  $Y$  - عدد المباني التي تم إنشاؤها في مدينة جدة .
- ج -  $Z$  - عدد البيضات التي تبيضها دجاجة خلال شهر أيلول .
- د -  $M$  - قيمة الحليب الذي تنتجه بقرة في مزارع فقيه للدواجن خلال عام .
- (٢) يحتوى صندوق على أربع كرات سود وكرتين خضريتين . سحبنا ثلاث كرات من الصندوق على التالى ، ومع إعادة الكرة المسحوبة إلى الصندوق قبل سحب الأخرى ، ما هو التوزيع الاحتمالى لعدد الكرات الخضريتين المسحوبة من هذا الصندوق ؟
- (٣) أوجد عبارة التوزيع الاحتمالى للمتغير العشوائى  $X$  الممثل للنتيجة التى تظهر عند إلقاء حجر نرد مرة واحدة .
- (٤) بضاعة مصنعة مؤلفة من ست قطع من بينها قطعتان معابتان ، أوجد التوزيع الاحتمالى للمتغير  $X$  الممثل لعدد القطع المعيبة لدى سحب ثلاث قطع من هذه البضاعة ، وذلك بصورة عشوائية . عبر عن النتائج بوساطة ما يسمى بمخطط توزيع التواتر .
- (٥) أوجد فى المثال ٤ التوزيع التراكمى للمتغير  $X$  . استخدام  $F(x)$  للبحث عن :
- أ —  $P[X = 1]$
- ب —  $P[0 < X < 2]$
- (٦) إذا علمت أن الكثافة الاحتمالية لمتغير عشوائى مستمر  $X$  هي من الشكل :

$$f(x) = \begin{cases} k \sqrt{x} & : 0 < x < 1 \\ 0 & : \text{فيما عدا ذلك} \end{cases}$$

فالمطلوب :

أ — تحديد قيمة الثابت K

ب — إيجاد  $F(x)$  ، ثم استخدمه لحساب قيمة  $P(0.3 < X < 0.6)$  .

(٧) يحتوى صندوق فاكهة على ثلاث برتقالات وتفاحتين وثلاث موزات . سحبنا وبصورة عشوائية أربع قطع من الصندوق . فإذا كان  $X$  ممثلاً لعدد البرتقالات المسحوبة ،  $Y$  ممثلاً لعدد التفاحات المسحوبة أيضاً فالمطلوب :

أ — إيجاد التوزيع الاحتمالي المشترك للمتغيرين  $(X, Y)$  .

ب —  $P[(x, y) \in A]$  حيث تمثل  $A$  المنطقة من المستوى المحددة بالجموعة  $\{(x, y) | x + y \leq 2\}$

(٨) متغيران عشوائيان مستمران لهما كثافة احتمالية مشتركة محددة بالعلاقة :

$$f(x, y) = \begin{cases} 4xy & : 0 < x < 1, 0 < y < 1 \\ 0 & : \text{فيما عدا ذلك} \end{cases}$$

والمطلوب :

أ — ما هو  $P[0 \leq x < \frac{3}{4} | \frac{1}{8} \leq y \leq \frac{1}{2}]$

ب — ما هو  $P[Y > X]$

(٩) متغيران عشوائيان مستمران لهما كثافة احتمالية مشتركة من الشكل :

$$f(x, y) = \begin{cases} K(x^2 + y^2) & : 0 < x < 2, 1 < y < 4 \\ 0 & : \text{فيما عدا ذلك} \end{cases}$$

فالمطلوب :

- أ — تحديد قيمة  $K$   
 ب — إيجاد قيمة الاحتمال  
 ج — ما هي قيمة  
 د — إيجاد قيمة الاحتمال
- $P[1 < X < 2, 2 < Y \leq 3]$   
 $P[1 \leq X \leq 2]$   
 $P[X + Y > 4]$

(١٠) متغيران عشوائيان بالكثافة المشتركة :

$$f(x,y) = \begin{cases} \frac{1}{y} & : 0 < x < y, 0 < y < 1 \\ 0 & \text{فيما عدا ذلك} \end{cases}$$

أوجد :  $P[X + Y > \frac{1}{2}]$ 

(١١) عد إلى التمرين (٥) واحسب :

- أ —  $f(y/2)$   
 ب —  $P[Y = 0 | X = 2]$

(١٢) إذا علمت أن التوزيع الاحتمالي المشترك للمتغيرين العشوائيين  $X, Y$  معطى بالجدول التالي :

Y \ X	X		
	1	2	3
1	0	$\frac{1}{6}$	$\frac{1}{12}$
2	$\frac{1}{5}$	$\frac{1}{9}$	0
3	$\frac{2}{15}$	$\frac{1}{4}$	$\frac{1}{15}$

فأوجد التوزيعات الهامشية والشرطية لهذين المتغيرين .

(١٣) بفرض أن الكثافة المشتركة للمتغيرين المستمرين  $(X, Y)$  هي من الشكل :

$$f(x, y) = \begin{cases} \frac{6-x-y}{8} & : 0 < x < 2, \\ & : 2 < y < 4 \\ 0 & : \text{فيما عدا ذلك} \end{cases}$$

أوجد  $P\{1 < Y < 3 | X = 2\}$

(١٤) بفرض أن الكثافة الاحتمالية المشتركة للمتغيرين  $(X, Y)$  هي من الشكل :

$$f(x, y) = \begin{cases} 2 & : 0 < x < y < 1 \\ 0 & : \text{فيما عدا ذلك} \end{cases}$$

فالمطلوب :

أ — تحديد فيما إذا كان المتغيران مستقلين .

ب — إيجاد :

$$P\left\{\frac{1}{4} < X < \frac{1}{2} \mid Y = \frac{3}{4}\right\}$$

(١٥) متغير عشوائى توزيعه الاحتمالى معطى بالجدول التالى :

X	-3	6	9
P [X = x]	$\frac{1}{6}$	$\frac{1}{2}$	$\frac{1}{3}$

والمطلوب :

أ — حساب قيمة  $E(X^2), E(X)$

ب — استخدام قوانين التوقع الرياضى فى حساب :

$$E[(2X + 1)^2], E[\{X - E(X)^2\}]$$

(١٦) بفرض أن  $X$  يمثل العدد الذى ظهر لدى إلقاء حجر نرد أحمر ،  $Y$  العدد الذى ظهر لدى إلقاء نرد أسود . أوجد :

أ —  $E(X + Y)$

ب —  $E(X - Y)$

ج —  $E(X \cdot Y)$

(١٧) أوجد تمام تباين المتغيرين العشوائيين فى التمرين (٧) .

(١٨) أوجد تمام المتغيرين العشوائيين فى التمرين (١٢) .

(١٩) برهن أن :  $\text{Cov}(ax, by) = a \cdot b \text{Cov}(x, y)$

(٢٠) متغير عشوائى  $X$  له متوسط  $\mu = 12$  وتباين  $\sigma^2 = 9$  ، وتوزيعه الاحتمالى غير معلوم . مستخدماً متباينة تشيشفيف أوجد :

أ —  $P[6 < X < 18]$

ب —  $P[3 < X < 21]$

(٢١) برهن متباينة تشيشفيف من أجل أى متغير عشوائى  $X$  منقطع .

(٢٢) متغير عشوائى مستمر دالة كثافته من الشكل :

$$f(x) = \begin{cases} \frac{1}{2} \cos x & : x \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \\ 0 & : x \notin \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \end{cases}$$

احسب توقع وتباين هذا المتغير .



## الفصل الثالث

### بعض توزيعات الاحتمال المنقطعة

- مقدمة ■ التوزيع المنتظم ■ التوزيع الحداني والتوزيع المتعدد الحدود
- التوزيع الهندسي الزائدي ■ التوزيع البواسوني ■ التوزيع الحداني
- السالب ■ ثمارين معلولة ■ ثمارين عامة .



### (٣,١) مقدمة

لقد أوضحنا في الفصل الثاني أن توزيع الاحتمال المنقطع يمثل بمجدول يوضح قيم المتغير العشوائى ، والاحتمالات الموافقة لهذه القيم . والمسألة المطروحة الآن ما هى الطريقة التى تصف لنا سلوك متغير عشوائى ما . الملاحظ أن بعض المتغيرات العشوائية والتى تتعلق ببعض التجارب الإحصائية تنصف بخواص متشابهة ، ويمكن أن نصف بنفس توزيع الاحتمال ، فمثلا أن جميع المتغيرات العشوائية الممثلة لعدد مرات النجاح فى  $n$  اختبارا معادا لتجارب مستقلة ( حيث يمثل احتمال الحصول على نجاح فى كل اختبار مقدارا ثابتا لا يتغير من اختبار لآخر ) يتصف سلوكها بنفس النموذج ، ولذلك يمكن تمثيلها بصيغة خاصة .

فى هذا الفصل سندرس عددا هاما من توزيعات الاحتمال المنقطعة . إن هذه التوزيعات تصف لنا معظم المتغيرات العشوائية التى تصادفنا خلال تجاربنا العملية .

### (٣,٢) التوزيع المنتظم Uniform distribution

إن من أبسط توزيعات الاحتمال المنقطعة هو ذاك التوزيع الذى يفترض فيه المتغير العشوائى جميع قيمه باحتمالات متساوية . مثل هذا التوزيع يدعى بالتوزيع المنتظم .

#### التوزيع المنتظم

إذا افترض المتغير العشوائى  $X$  جميع القيم  $x_1, x_2, \dots, x_n$  باحتمالات متساوية ، فإننا نعرف عندئذ توزيع الاحتمال لهذا المتغير المنتظم بالعلاقة :

$$f(x; n) = \frac{1}{n} : x = x_1, x_2, \dots, x_n$$

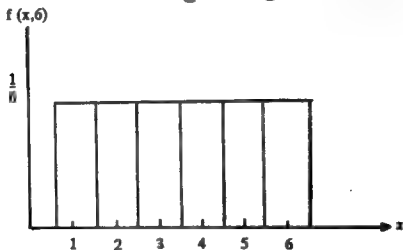
لقد استخدمنا الرمز  $f(x; n)$  بدلا عن  $f(x)$  لنشير إلى أن التوزيع المنتظم يتعلق بـ  $n$  (عدد القيم التي يفترضها المتغير العشوائي) .

### مثال (٣, ١)

عند إلقاء حجر نرد متوازن ومتمائل فإننا نجد أن كل نتيجة في فضاء العينة  $S = \{1, 2, \dots, 6\}$  يمكن أن تظهر باحتمال قدره  $\frac{1}{6}$  لذلك فإن المتغير العشوائي الممثل للعدد الذي سيظهر يمثل متغيراً عشوائياً منتظماً .

### الحل

إن المنحنى البياني الممثل لتوزيع المنتظم يمثل مجموعة مستطيلات بارتفاعات متساوية . والشكل (٣, ١) يوضح هذا التوزيع .



الشكل (٣, ١)

### نظرية (٣, ١)

توقع وتباين متغير عشوائي منتظم توزيعه  $f(x; n)$  يعطى بالعلاقين التاليين :

$$\mu = \frac{\sum_{i=1}^n x_i}{n} , \quad \sigma^2 = \frac{\sum_{i=1}^n (x_i - \mu)^2}{n}$$

## الرهان

من تعريف التوقع الرياضى لمتغير عشوائى نكتب :

$$\begin{aligned}\mu = E(X) &= \sum_{i=1}^n x_i \cdot f(x_i; n) = \sum_{i=1}^n \frac{x_i}{n} \\ &= \frac{\sum_{i=1}^n x_i}{n}\end{aligned}$$

وحسب تعريف التباين نجد أن :

$$\begin{aligned}\sigma^2 = E[(X - \mu)^2] &= \sum_{i=1}^n (x_i - \mu)^2 f(x_i; n) \\ &= \sum_{i=1}^n \frac{(x_i - \mu)^2}{n} = \frac{\sum_{i=1}^n (x_i - \mu)^2}{n}\end{aligned}$$

مثال (٣، ٢)

احسب فى المثال (٣، ١) توقع وتباين المتغير الممثل للوجه الذى ظهر لدى إلقاء قطعة زهر .

## الحل

حسب النظرية (٣، ١) لدينا :

$$\begin{aligned}\mu &= \frac{\sum_{i=1}^n x_i}{n} = \frac{1+2+3+4+5+6}{6} = 3.5 \\ \sigma^2 &= \frac{(1-3.5)^2 + (2-3.5)^2 + \dots + (6-3.5)^2}{6} = \frac{35}{12}\end{aligned}$$

## (٣، ٣) التوزيع الحدائى والتوزيع المتعدد الحدود

## Binomial and multinomial distribution

يقترن أحد أهم التوزيعات العشوائية المنقطعة بتجربة إلقاء قطعة النقود التى ذكرناها فى الأمثلة (١، ٦) ، (١، ١٨) ، (١، ٢١) ، (١، ٢٦) ،...، وإلغ وتتكرر يومياً العديد من التجارب التى تشبه تجربة إلقاء النقود ذات الأهمية التطبيقية فى العلوم المختلفة .

فالرمي على هدف يشبه إلى حد كبير إلقاء قطعة النقود لأن عملية الرمي تقود إلى إحدى نتيجتين إما إصابة للهدف أو إخطاء له ، كذلك الأمر في فحص فعالية دواء جديد فإذا ما أن يكون هذا الدواء فعالاً أو غير فعال ، كذلك عند معرفة رأى ناخب في مرشح ما ، إما أن يصوت ضد أو مع هذا المرشح ، وأخيراً فإن فحص قطعة من بضاعة مصنعة سيرتك لنا الباب مفتوحاً للحكم على نوع هذه القطعة إن كانت معيبة أو جيدة الصنع . كل هذه الأمثلة وغيرها تكشف لنا أن هذه التجارب تتشابه إلى حد مقبول في الخواص التالية :

- (١) تتألف كل تجربة من عدد  $n$  من الاختبارات المتماثلة تماماً .
- (٢) كل تكرار للتجربة ينتج عنه إحدى نتيجتين إما نجاح أو فشل . فسؤالنا عن فعالية دواء مثلاً ضد مرض معين سيتحدد بإحدى النتيجتين ، فعال ( نجاح ) أو غير فعال ( فشل ) كذلك إصابة رام للهدف يمثل نجاحاً ، وإخطاؤه للهدف فشلاً .
- (٣) إن احتمال النجاح في كل باختبار يبقى ثابتاً . فمثلاً يصيب رامى الهدف باحتمال قدره 0.7 ، وهذا الاحتمال لا يتغير سواء في الرمية الأولى أو العاشرة . كذلك فإن احتمال مقابلة ناخب مؤيد للمرشح الفلاني ، يبقى ثابتاً تقريباً طالما كان يجمع الناخبين كبيراً جداً بالمقارنة مع العينة من الناخبين الذين تجرى مقابلتهم . فإذا كان 50% مثلاً من المجتمع الأمريكى يحوى ألف ناخب يفضلون المرشح إدوارد كنيدي ، فإن احتمال الحصول على تأييد لهذا المرشح عند أول مقابلة هو  $\frac{1}{2}$  . واحتمال التأييد عند المقابلة الثانية هو  $\frac{500}{999}$  أو  $\frac{499}{999}$  حسباً تكون المقابلة الأولى قد تمت مع مؤيد أو مع معارض على الترتيب . والعديدان قريبان جداً من  $\frac{1}{2}$  ويمكن اعتبارهما  $\frac{1}{2}$  عملياً ، كما يمكن أن نستمر في مثل هذا والرابعة حتى المقابلة الـ  $n$  طالما ظلت  $n$  صغيرة بالنسبة للعدد 1000 ، وبصورة خاصة إذا كان عدد الناخبين عشرة ، وكان خمسة منهم يفضلون المرشح إدوارد كنيدي ، فإن احتمال الحصول على تأييد في المقابلة الأولى هو  $\frac{1}{2}$  ، أما في الثانية فهو  $\frac{4}{9}$  أو  $\frac{5}{9}$  . أى أن الاحتمال  $P$  يتغير كثيراً من اختبار إلى آخر ، والتجربة لن تكون تجربة حدانية .
- (٤) الاختبارات المكررة مستقلة .

وبذلك تتمتع التجربة الحدانية بالخواص الأربعة التالية :

- (١) تتألف كل تجربة من  $n$  اختبار معاداً .

- (٢) إن نتيجة كل اختبار إما نجاح أو فشل .  
 (٣) احتمال النجاح في كل اختبار ثابت لا يتغير من اختبار لآخر .  
 (٤) الاختبارات المكررة مستقلة .

لنفرض على سبيل المثال أن تجربتنا تنحصر في سحب عنصرين من مجموعة بضاعة مصنعة بشكل عشوائي . سنرمز بـ  $N$  للعنصر المسحوب من نوع جيد و بـ  $D$  للعنصر المعاب ، ومنفرض أن العنصر المعاب هو بمثابة نجاح . أن عدد النجاحات  $X$  هو متغير عشوائي يفترض القيم 0, 1, 2, وفي هذه التجربة نجد النتائج الأربع التالية :

النتيجة	X
NN	0
DN	1
ND	1
DD	2

لنفرض أن عدد العناصر المعابة هو 35% . بما أن العناصر المسحوبة قد سحبت بشكل عشوائي لذلك فإننا نجد :

$$P(ND) = P(N) \cdot P(D) = (0.65) (0.35) = 0.2275$$

وبشكل مشابه نحسب احتمالات بقية النقاط في الجدول السابق فنجد أن :

X	0	1	2
P	0.4225	0.455	0.1225

**تعريف (٣,١) متغير عشوائي حداثي Binomial Random Variable**

إن عدد مرات النجاح  $X$  في تجربة حداثية ( مكونة من  $n$  اختبارا مكررا ) يدعى بمتغير عشوائي حداثي .

نسمى توزيع الاحتمال للمتغير الحداثي بالتوزيع الحداثي ، وسنرمز له بالرمز

$b(x; n, p)$  ، لأن قيم هذا الاحتمال تتعلق بعدد الاختبارات المكررة و باحتمال النجاح في كل اختبار . مثلاً في المثال السابق نجد أن :

$$P(X = 2) = f(2) = b(2; 3, 0.35) = 0.1225$$

لنعمم الآن المناقشة السابقة لنحصل على صيغة للتوزيع  $b(x; n, p)$  . إن ما نرغب به هو الحصول على صيغة تعطينا احتمال الحصول على  $x$  نجاح عند تكرار تجربة حدائية  $n$  مرة . لنحسب أولاً احتمال الحصول على  $x$  نجاح و  $n - x$  فشل في ترتيب معين . بما أن الاختبارات المكررة مستقلة ، لذلك يمكن أن نحصل على الصيغة المطلوبة من حاصل ضرب كل الاحتمالات الموافقة للنتائج المختلفة . فإذا رمزنا لاحتقال النجاح بالرمز  $P$  و لاحتقال الفشل بالرمز  $q = 1 - p$  فعندئذ يكون الاحتمال المطلوب في ترتيب معين مساوياً  $p^x \cdot q^{n-x}$  . علينا الآن أن نحدد المجموع الكلي لعدد نقاط العينة في تجربة حدائية فيها  $x$  نجاح و  $n - x$  فشل ، وهذا العدد هو :  $\left( \frac{n}{x} \right)$  لذلك فاحتمال الحصول على  $x$  نجاح في التجربة الحدائية هو  $p^x \cdot q^{n-x} \cdot \left( \frac{n}{x} \right)$  .

### التوزيع الحدائي

إذا كان احتمال النجاح في تجربة حدائية  $p$  ، واحتمال الفشل  $q = 1 - p$  ، عندئذ يكون توزيع الاحتمال للمتغير العشوائي الحدائي الممثل لعدد مرات النجاح في  $n$  اختباراً مستقلاً هو :

$$b(x; n, p) = \left( \frac{n}{x} \right) p^x \cdot q^{n-x} : x = 0, 1, 2, \dots, n$$

لاحظنا أنه إذا كان  $n = 2$  ،  $p = 0.35$  . فعندئذ يكون توزيع الاحتمال للمتغير  $X$  الممثل لعدد العناصر المعيبة عند سحب عنصرين بشكل عشوائي من بضاعة مصنعة هو :

$$b(x; 2, 0.35) = \left( \frac{2}{x} \right) \cdot (0.35)^x (0.65)^{2-x} : x = 0, 1, 2,$$

وهي نفس الصيغة المستخدمة في الجدول السابق .



## مثال (٣.٣)

بفرض أن احتمال أن بقاء مريض على قيد الحياة بعد إجراء عملية جراحية معينة هو 0.85 ، ما هو احتمال أن يبقى ثلاثة من خمسة مرضى آخرين على قيد الحياة عند إجراء نفس العملية لهم ؟

## الحل

نلاحظ أن إجراء عملية لمريض يمثل اختباراً مستقلاً عن نتيجة أى عملية أخرى تجرى لمريض آخر . ثم إن  $p = 0.85$  لكل اختبار ، لذلك :

$$\begin{aligned} b(3; 5, 0.85) &= \binom{5}{3} (0.85)^3 (0.15)^2 \\ &= \frac{5!}{3!2!} (0.85)^3 (0.15)^2 \\ &= 0.138 \end{aligned}$$

## ملاحظة

نلاحظ أن توزيع الاحتمال الحداني يستمد اسمه من حقيقة أن الـ  $n + 1$  عنصراً في منشور نيوتن  $(p + q)^n$  والتي توافق القيم  $x = 0, 1, \dots, n$  .

$$\begin{aligned} (q + p)^n &= \binom{n}{0} q^n + \binom{n}{1} p \cdot q^{n-1} + \binom{n}{2} p^2 \cdot q^{n-2} + \dots + \binom{n}{n} p^n \\ &= b(0; n, p) + b(1; n, p) + b(2; n, p) \dots + b(n; n, p) \end{aligned}$$

وبما أن  $P + q = 1$  لذلك نلاحظ أن :  $\sum_{x=0}^n b(x; n, p) = 1$  وهى الخاصية التى يجب أن يحققها أى توزيع احتمال .

من الملاحظ أيضاً أننا نهم على الغالب فى حساب  $P[X < r]$  ، أو  $P[a \leq x \leq b]$  . ولحسن الحظ فإن جدول II يعطى قيم المجموع  $\sum_{x=0}^n b(x; n, p) = B(r; n, p)$  ، وذلك من أجل  $n = 5, 10, 15, 20$  ، ومن أجل أى قيمة لـ  $p$  من 0.1 إلى 0.9 .

## مثال (٣، ٤)

أطلق صياد خمس طلقات على هدف . فإذا علمت أن احتمال إصابته للهدف في كل إطلاق هو 0.8 فما هو احتمال :

- أ — أن يصيب الهدف مرتين تماما ؟  
 ب — أن يصيب الهدف مرتين على الأقل ؟  
 ج — أن يصيب الهدف خمس مرات تماما ؟

## الحل

نلاحظ أن عملية الإطلاق على الهدف تتم على شكل تكرارات ( اختبارات ) مستقلة بعضها عن بعض ، كما نلاحظ أن عملية الإطلاق هذه تمثل تجربة حدانية فيها  $q = 0.2, p = 0.8, n = 5$  . بفرض أن عدد مرات إصابة الهدف هو  $X$  فإننا نجد أن :

$$P [ X = 2 ] = b(2; 5, 0.8)$$

$$= \sum_{x=0}^2 b(x; 5, 0.8) - \sum_{x=0}^1 b(x; 5, 0.8)$$

وبالعودة إلى جدول II في نهاية الكتاب نجد أنه من أجل  $p = 0.8, r = 2, n = 5$

$$\sum_{x=0}^2 b(x; 5, 0.8) = 0.0579$$

كذلك فإنه من أجل  $p = 0.8, r = 1, n = 5$  :

$$\sum_{x=0}^1 b(x; 5, 0.8) = 0.0067$$

وحاصل الطرح بين القيمتين هو :

$$p [ X = 2 ] = 0.0579 - 0.0067 \\ = 0.0512$$

ولو حسبنا هذا الاحتمال بالطريقة العادية لوجدنا أن :

$$\begin{aligned} P [X = 2] &= \binom{5}{2} (0.8)^2 \cdot (0.2)^3 \\ &= \frac{5!}{2!3!} (0.8)^2 (0.2)^3 \\ &= 0.0512 \end{aligned}$$

وهي نفس النتيجة السابقة .  
كذلك فإن :

$$\begin{aligned} P [X \geq 2] &= 1 - P [X < 2] \\ &= 1 - \sum_{x=0}^1 b(x; 5, 0.8) \end{aligned}$$

ومن الجدول II نجد أن المجموع السابق يساوى :

$$\begin{aligned} &= 1 - 0.0067 \\ &= 0.9933 \end{aligned}$$

وأخيرا :

$$\begin{aligned} P [X = 5] &= \sum_{x=0}^5 b(x; 5, 0.8) - \sum_{x=0}^4 b(x; 5, 0.8) \\ &= 1 - 0.6723 \\ &= 0.3277 \end{aligned}$$

ونلاحظ أيضا أن :

$$\begin{aligned} P [X = 5] &= \binom{5}{5} (0.8)^5 (0.2)^0 \\ &= (1) (0.8)^5 (1) \end{aligned}$$

$$= 0.32768$$

وهى تقريبا نفس النتيجة السابقة .

### مثال (٣,٥)

بفرض أن احتمال شفاء مريض في عملية جراحية هو 0.9 . فما هو احتمال أن يشفى ثمانية مرضى من أصل عشرة أجريت لهم نفس العملية ؟

### الحل

نلاحظ أن العمليات العشرة التى أجريت تمثل عشرة اختبارات في تجربة حدانية فيها نتيجة كل اختبار ، إما شفاء للمريض أو عدمه ، كما نلاحظ أن احتمال شفاء المريض في كل عملية ( اختبار ) ثابت من عملية إلى أخرى . لذلك فالتجربة المجراة هى تجربة حدانية فيها  $x = 8, p = 0.9, n = 10$  ، وذلك بفرض أن عدد المرضى الذين تم شفاؤهم هو  $X$  . أما الاحتمال المطلوب فهو :

$$P[X = 8] = P[X \leq 8] - P[X \leq 7]$$

$$= \sum_{x=0}^8 b(x; 10, 0.9) - \sum_{x=0}^7 b(x; 10, 0.9)$$

وبالعودة إلى جدول II نجد أن :

$$P[X = 8] = 0.2639 - 0.0702 \\ = 0.1937$$

### نظرية (٣,٢)

إن توقع وتباين متغير عشوائى حدانى  $X$  يعطيان بالعلاقين :

$$\mu = np, \sigma^2 = np \cdot q$$

### البرهان

لنرمز بـ  $X_i$  لعدد مرات النجاح في الاختبار رقم  $i$  للتجربة الحدانية . نلاحظ أن

المتغيرات  $X_i$  مستقلة ، كما أن كل متغير من هذه المتغيرات يفترض إحدى القيمتين 0 ، 1 ،  
 حسبما تكون نتيجة الاختبار فشل أو نجاح على الترتيب . كما نلاحظ أيضا أن مجموع  
 عدد مرات النجاح هو :

$$X = X_1 + X_2 + \dots + X_n$$

والملاحظ أن التوقع الرياضي لكل متغير من المتغيرات  $X_i$  هو :

$$E(X_i) = 1.p + 0.q = p$$

ولذلك مهما تكن  $i = 1, 2, \dots, n$  . وحسب النظرية (٣،٣) نجد أن :

$$\begin{aligned} E(X) &= E(X_1) + E(X_2) + \dots + E(X_n) \\ &= p + p + p + \dots + p \\ &= np \end{aligned}$$

أما تبين المتغير الحداني  $X$  فيحسب كما يلي :

$$\sigma_x^2 = \sigma_{x_1 + x_2 + \dots + x_n}^2$$

وحسب النظرية (١١،٢) وعلى اعتبار أن المتغيرات  $X_1, \dots, X_n$  مستقلة فإننا نجد أن :

$$\sigma_x^2 = \sigma_{x_1 + x_2 + x_3 + \dots + x_n}^2$$

غير أن :

$$\begin{aligned} \sigma_{x_1}^2 &= E(X_1^2) - p^2 \\ &= (0)^2 \cdot q + (1)^2 \cdot p - p^2 \end{aligned}$$

$$\sigma_{x_1}^2 = p.q$$

$$\sigma_x^2 = n p.q$$

وهو المطلوب برهانه

مثال (٣،٦)

احسب توقع وتباين عدد المرضى الذين تم شفاؤهم في المثال (٣،٥) .

## الحل

نلاحظ أن :

$$n = 10$$

$$p = 0.9$$

$$q = 0.1$$

باستخدام النظرية (٣، ٢) نجد أن :

$$\mu = n.p = (10)(0.9) = 9$$

$$\sigma^2 = n.p.q = (10)(0.9)(0.1) = 0.9$$

لنفرض الآن لكل اختبار ( في التجربة الحدانية ) أكثر من نتيجتين ممكنتين ، فعندئذ نقول بأن لدينا تجربة متعددة الحدود (multinomial experiment) . فلو اعتبرنا أن ورق اللعب مؤلف من أربعة أنواع بحسب النوع ( دينارى — بستوى — كبه — سباقى ) فعند سحب ورقة من ورق اللعب مع الإعادة بصورة عشوائية نكون أمام تجربة متعددة الحدود ، لأن نتيجة السحب ستكون واحدة من أربعة أنواع .

وبشكل عام ، إذا كانت نتيجة التجربة إحدى النتائج الـ  $k$  الممكنة  $E_1, E_2, \dots, E_k$  ، وإذا كانت الاحتمالات الموافقة لهذه النتائج  $P_1, \dots, P_k$  ، فعندئذ يمثل التوزيع المتعدد الحدود احتمال الحصول على النتائج  $E_1, E_2, \dots, E_k$  عدداً من المرات  $X_1, \dots, X_k$  على الترتيب في  $n$  تكراراً مستقلاً ، حيث إن  $X_1, X_2, \dots, X_n = n$  .

سنرمز لتوزيع الاحتمال هذا بالرمز  $f(x_1, \dots, x_k, P_1, \dots, P_k, n)$  من الواضح أن  $P_1 + \dots + P_k = 1$  ، لأن نتيجة كل اختبار ( أو تكرار ) تمثل إحدى النتائج الـ  $k$  الممكنة .

وللحصول على الصيغة العامة لهذا التوزيع ، ستتع نفس الخطوات التى سرنا عليها لدى الحصول على التوزيع الحداني . نفرض ترتيب معين للنتائج ، ثم نحسب احتمال هذا الترتيب فنجد أنه :  $P_1^{x_1} P_2^{x_2} \dots P_k^{x_k}$

ثم نبحث عن المجموع العام لهذه الترتيبات ، وهذا العدد يمثل عدد الطرق التى يمكن أن

نجزأ بها مجموعة مؤلفة من  $n$  عنصراً إلى عدد من الخلايا  $k$  تحوى الأولى على  $x_1$  عنصراً ،  
والثانية على  $x_k$  عنصراً . وهذا العدد من التجزئات يمكن أن يتم بعدد من الطرق مساو  
لـ :

$$\left( \begin{matrix} n \\ x_1, x_2, \dots, x_k \end{matrix} \right) = \frac{n!}{x_1! x_2! \dots x_k!}$$

وبما أن كل التجزئات متنافية تبادلياً ( أى لا يمكن أن نحصل على تجزئتين مختلفتين في  
وقت واحد ) ، واحتمال وقوع أى منها واحد ويساوى  $P_1^{x_1} \dots P_k^{x_k}$  ، لذلك فإن  
الاحتمال المطلوب هو :

$$f(x_1, \dots, x_k; p_1, \dots, p_k, n) = \frac{n!}{x_1! \dots x_k!} p_1^{x_1} \dots p_k^{x_k}$$

#### التوزيع المتعدد الحدود Multinomial distribution

إذا كانت النتائج الممكنة لتجربة معينة هي  $E_1, E_2, \dots, E_k$  ، واحتمالاتها الموافقة  
هي  $p_1, \dots, p_k$  فنعتقد يكون توزيع الاحتمال للمتغيرات  $X_1, \dots, X_k$  الممثلة لعدد  
المرات التى متفع فيها النتائج  $E_1, \dots, E_k$  على الترتيب في  $n$  تجربة مستقلة معطى بالعلاقة  
السابقة .

$$\sum_{i=1}^k x_i = n, \sum_{i=1}^k p_i = 1 \quad \text{وحيث إن :}$$

ومما يجدر ذكره أن التوزيع المتعدد قد اشتق اسمه من حقيقة أن المنشور المتعدد  
هو  $f(x_1, \dots, x_k; p_1, \dots, p_k, n)$  يوافق جميع قيم  $(p_1 + p_2 + \dots + p_k)^n$  .

#### مثال (٣،٧)

لنحسب احتمال الحصول على مجموع يساوى 3 أو 9 برتين ، ثنائية متشابهة مرة  
واحدة ، وأى شكل آخر ثلاث مرات لدى إلقاء حجرى نرد ست مرات متتالية .

الحل :

لنشكل الحوادث التالية :

$A_1 = \{ \text{الحصول على مجموع 3 أو 9} \}$

$A_2 = \{ \text{الحصول على ثنائية متشابهة} \}$

$A_3 = \{ \text{الحصول على أى شكل آخر} \}$

والمراد الآن أن تقع الحوادث  $A_1$  مرتين ،  $A_2$  مرة واحدة ،  $A_3$  ثلاث مرات .

نلاحظ أن :

$$P_1 = P(A_1) = \frac{4+2}{36} = \frac{1}{36}, P_2 = P(A_2) = \frac{6}{36} = \frac{1}{6}$$

$$P_3 = P(A_3) = 1 - \left[ \frac{1}{36} + \frac{6}{36} \right] = 1 - \frac{7}{36} = \frac{29}{36}$$

لذلك فالاحتمال المطلوب هو :

$$f(2, 1, 3; \frac{1}{36}, \frac{1}{6}, \frac{29}{36}, 6) = \frac{6!}{2! \cdot 1! \cdot 3!} \left( \frac{1}{36} \right)^2 \cdot \left( \frac{1}{6} \right)^1 \cdot \left( \frac{29}{36} \right)^3$$

### (٣,٤) التوزيع الهندسى الزائدى Hypergeometric distribution

من الملاحظ أنه لا يمكن تطبيق التوزيع الحداني في حساب احتمال الحصول على أربع أوراق حمراء ، وذلك عند سحب ست أوراق من ورق اللعب بصورة عشوائية دفعة واحدة أو بالتالي ولكن بدون إعادة إلا أنه إذا سحبنا ورقة بصورة عشوائية وأعدناها الى مجموعة ورق اللعب ثم خلطنا ورق اللعب خلطاً جيداً ، وكررنا العملية السابقة مثلاً ست مرات ، فإنه يمكن عندئذ حساب الاحتمال السابق المذكور . بشكل عام يمكن حساب الاحتمال السابق عند السحب بدون إعادة على النحو التالى :

نفرض الحوادث :

$A = \{ \text{ظهور أربع أوراق حمراء} \}$

$B = \{ \text{ظهور ورقتين سوداوين} \}$



نلاحظ أن فضاء العينة S في هذه التجربة يتألف من  $\binom{52}{6}$  نقطة ، وبين هذه النقاط عددا يوافق الحادث A هو  $\binom{26}{4}$  . كما أنه يوافق الحادث B عددا من النقاط مساوٍ لـ  $\binom{26}{2}$  ، والاحتمال المطلوب يعطى بالعلاقة :

$$P = \frac{\binom{26}{4} \binom{26}{2}}{\binom{52}{6}} = 0.23$$

نسمى تجربة سحب n شيئا في مجموعة من الأشياء ( مؤلفة من N شيئا منها K شيئا مكتوب على كل منها نجاح و N - K فشلا ) وظهور x نجاحا و n - x فشلا بتجربة هندسية زائدة . من الملاحظ أن التجربة الهندسية الزائدة تتصف بالخواص التالية :

- (١) سحب n عنصرا من N شيئا .
- (٢) السحب يتم دفعة واحدة وبدون إعادة .
- (٣) K شيئا من الـ N المفروضة مكتوب على كل منها نجاح و N - K فشل .

### تعريف (٣،٢) التوزيع الهندسي الزائدي

نسمى عدد النجاحات في تجربة هندسية زائدة بالمتغير الهندسي الزائدي ، كما أن احتمال هذا المتغير يدعى بالتوزيع الهندسي الزائد ، ونظرا لكون هذا الاحتمال معتمدا على K, N فإننا سنرمز له بالرمز  $h(x; N, n, k)$  .

### مثال (٣،٨)

اختيرت لجنة مؤلفة من أربعة أشخاص بشكل عشوائي من مجموعة أربعة كيميائيين ، وسبعة فيزيائيين . ما هو توزيع الاحتمال لعدد الكيميائيين في اللجنة ؟

### الحل

نلاحظ أن التجربة المجراة هي تجربة هندسية زائدة . ل نرمز بـ X لعدد الكيميائيين الموجودين في اللجنة . نجد أن قيم المتغير X هي 0, 1, 2, 3, 4 ، وتوزيع الاحتمال للمتغير الهندسي الزائدي x هو :

$$h(x; 11, 4, 7) = P\{X = x\}$$

كملا نلاحظ أن :

$$h(0; 11, 4, 7) = \frac{\binom{4}{0} \binom{7}{4}}{\binom{11}{4}} = 0.106$$

$$h(1; 11, 4, 7) = \frac{\binom{4}{1} \binom{7}{3}}{\binom{11}{4}} = 0.424$$

$$h(2; 11, 4, 7) = \frac{\binom{4}{2} \binom{7}{2}}{\binom{11}{4}} = 0.38$$

$$h(3; 11, 4, 7) = \frac{\binom{4}{3} \binom{7}{1}}{\binom{11}{4}} = 0.0848$$

$$h(4; 11, 4, 7) = \frac{\binom{4}{4} \binom{7}{0}}{\binom{11}{4}} = 0.0030$$

وهكذا نجد أن جدول توزيع X هو :

X	0	1	2	3	4
P	0.106	0.424	0.38	0.0848	0.00303

ومن السهل ملاحظة أن توزيع الاحتمال يمكن الحصول عليه من الصيغة التالية :

$$h(x; 11, 4, 7) = \frac{\binom{4}{x} \cdot \binom{7}{4-x}}{\binom{11}{4}} : x=1, 0, 2, 3, 4$$

لنعمم الآن المثال (٣,٨) في البحث عن احتمال سحب  $n$  عنصرا من ضمن  $N$  شيئا وظهور  $x$  نجاح ، و  $n-x$  فشل . أن عدد إمكانيات اختيار  $n$  شيئا ( في المجموعة التي تحتوي  $N$  عنصرا ) وبشكل عشوائي هو  $\binom{N}{n}$  طريقة . هذا العدد يمثل عدد نقاط فضاء العينة  $S$  . سنفرض أن هذه النقاط متساوية في إمكانية وقوعها .

نلاحظ أن هناك  $\binom{K}{x}$  طريقة لاختيار  $x$  نجاحا من ضمن  $K$  نجاح ، وكل طريقة من هذه الطرق توافق  $\binom{N-K}{n-x}$  طريقة لاختيار  $n-x$  فشلا من ضمن الـ  $N-K$  فشل ( والموجودة ضمن الـ  $N$  شيئا ) ، وهكذا نجد أنه يوافق الحادث المراد حساب احتماله عددا من النقاط  $\binom{N-K}{n-x} \cdot \binom{K}{x}$  ، إذن فالاحتمال المطلوب والذي نرمز له بالرمز  $h(x; N, n, k)$  يعطى من خلال التعريف التالي :

### التوزيع الهندسي الزائدي

إن التوزيع الاحتمالي للمتغير العشوائي الهندسي الزائدي  $X$  ( المثل لعدد مرات النجاح عند سحب  $n$  عنصرا بشكل عشوائي من  $N$  عنصرا تحتوي  $K$  نجاحا و  $N-K$  فشلا ) يعطى بالعلاقة التالية :

$$h(x; N, n, k) = \frac{\binom{k}{x} \cdot \binom{n-k}{n-x}}{\binom{N}{n}} , x = 0, 1, 2, \dots, n$$

نظرية (٣,٣)

إن توقع وتباين متغير عشوائي هندسي زائدي بالتوزيع  $h(x; N, n, k)$  يعطيان بالعلاقين التاليين :

$$\mu = \frac{nK}{N} , \sigma^2 = \frac{(N-n)}{(N-1)} \cdot n \cdot \frac{K}{N} \cdot (1 - \frac{K}{N})$$

### البرهان

للبحث عن التوقع الرياضي للمتغير الهندسي الزائدي نكتب ما يلي :

$$\begin{aligned}
 \mu = E(X) &= \sum_{x=0}^n x \frac{\binom{K}{x} \cdot \binom{N-K}{n-x}}{\binom{N}{n}} \\
 &= K \sum_{x=1}^n \frac{(K-1)!}{(x-1)!(K-x)!} \cdot \frac{\binom{N-K}{n-x}}{\binom{N}{n}} \\
 &= K \cdot \sum_{x=1}^n \frac{\binom{K-1}{x-1} \binom{N-K}{n-x}}{\binom{N}{n}} \quad Y = x-1
 \end{aligned}$$

وبوضوح  $Y = x - 1$  فإننا نجد أن :

$$\begin{aligned}
 \mu &= K \cdot \sum_{y=0}^{n-1} \frac{\binom{K-1}{y} \binom{N-K}{n-1-y}}{\binom{N}{n}} \\
 &= K \cdot \sum_{y=0}^{n-1} \frac{\binom{K-1}{y} \binom{(N-1)-(K-1)}{(n-1)-y}}{\binom{N}{n} \cdot \binom{N-1}{n-1}} \\
 &= \frac{Kn}{N}
 \end{aligned}$$

لأن المجموع الأخير يمثل مجموع كل الاحتمالات في التجربة الهندسية الزائدة ، عندما نختار  $(n-1)$  عنصرا من  $(N-1)$  عنصر نحوى على  $K-1$  نجاحا . وهذا المجموع يساوى الواحد حسب خواص الكثافة الاحتمالية .

لإيجاد تباين التوزيع الهندسي الزائدى ، نتبع نفس الخطوات السابقة للحصول على :

$$E[X(X-1)] = \frac{K \cdot (K-1) \cdot n \cdot (n-1)}{N \cdot (N-1)}$$

واعتاداً على النظرية (٧,٦) نجد أن :

$$\sigma^2 = E(X^2) - \mu^2$$

$$\begin{aligned} &= E[X(X-1)] + \mu - \mu^2 \\ &= \frac{K \cdot (K-1) \cdot n \cdot (n-1)}{N \cdot (N-1)} + \frac{nK}{N} - \left(\frac{nK}{N}\right)^2 \\ &= \frac{nK \cdot (N-K) \cdot (N-n)}{N^2 \cdot (N-1)} \\ &= \frac{(N-n)}{(N-1)} \cdot n \cdot \frac{K}{N} \cdot \left(1 - \frac{K}{N}\right) \end{aligned}$$

### مثال (٣,٩)

تحتوى كل رزمة من بضاعة مصنعة على 40 عنصراً ، وتقبل رزمة ما إذا حوت عدداً من العناصر المعيبة لا يتجاوز الثلاثة عناصر . لفحص أى رزمة نأخذ منها عينة مؤلفة من خمسة عناصر ونفحصها ، فإذا وجدنا فيها عنصراً معيباً فإننا نرفض الرزمة . ما هو احتمال وجود عنصر معيب فى عينة عند فحص رزمة . إذا علمت أن الرزمة المفحوصة تحتوى على ثلاثة عناصر معيبة ؟

### الحل

باستخدام التوزيع الهندسي الزائدى من أجل  $n = 5, N = 40, K = 3, x = 1$  ، فإننا نجد أن احتمال وجود عنصر معيب فى العينة هو :

$$h(1; 40, 5, 3) = \frac{\binom{3}{1} \binom{37}{4}}{\binom{40}{5}} = 0.3011$$

## مثال (٣, ١٠)

باستخدام نظرية تشبيشيف ، لنبحث عن المجال  $(\mu - 2\sigma, \mu + 2\sigma)$  في المثال السابق .

## الحل

اعتمادا على النظرية (٣, ٣) نجد أن :

$$\mu = \frac{(5) \cdot (3)}{40} = \frac{3}{8} = 0.375$$

$$\sigma^2 = \frac{(40-5)}{(39)} (5) \cdot \left(\frac{3}{40}\right) \cdot \left(1 - \frac{3}{40}\right) = 0.3113$$

وبأخذ الجذر التربيعي للمقدار الأخير ، فإننا نجد أن  $\sigma = 0.558$  ، والمجال المطلوب هو :

$$(-0.741, 1.491)$$

وحسب نظرية تشبيشيف نجد أن احتمال وقوع عدد العناصر المعيبة ( التي نجدها في عينة مؤلفة من خمسة عناصر مأخوذة من رزمة مؤلفة من 40 عنصرا فيها ثلاثة عناصر معيبة ) في المجال  $(-0.741, 1.491)$  على الأقل  $\frac{3}{4}$  . وهذا يعني أنه في ثلاثة أرباع الفترة الزمنية ستحتوى العينة المؤلفة من خمسة عناصر على عنصرين معييين على الأقل .

إذا كان  $n$  صغيرا بالنسبة لـ  $N$  فإننا نجد أن الاحتمال في كل سحبة سيتغير بشكل تافه . وهذا يعني أن تجربتنا ستأخذ شكل التجربة الحدانية وإذا يمكن تقريب التوزيع الهندسي الزائدي في هذه الحالة إلى التوزيع الحداني باستخدام  $P = \frac{K}{N}$  . ويقرب كل من المتوقع والتباين بوساطة العلاقات :

$$\mu = n.p. = \frac{n \cdot K}{N}$$

$$\sigma^2 = n.p.q. = n \cdot \frac{K}{N} \cdot \left(1 - \frac{K}{N}\right)$$

وبمقارنة هذه الصيغ مع ما جاء في النظرية (٣,٣) نلاحظ أن التوقع الرياضي ظل نفسه ، بينما اختلف التباين بتعديل العامل  $\frac{N-n}{N-1}$  ، هذا تافه عندما تكون  $n$  صغيرة بالنسبة لـ  $N$  .

### مثال (٣,١١)

أعلنت شركة لصنع الإطارات المطاطية أن شحنتها المؤلفة من 3200 إطار والمتجهة إلى أسواق البيع تحوى على 800 إطار معيب . لنحسب احتمال أن يشتري شخص ما عشرة إطارات من هذه الكمية بشكل عشوائى فيجد بينها إطارين معيبين .

### الحل

بما أن العدد  $N = 3200$  كبير بالنسبة لحجم العينة المختارة  $n = 10$  ، لذلك فإننا سنقوم بتقريب الاحتمال المطلوب باستخدام التوزيع الحدائى . أن احتمال الحصول على إطار معيب يساوى  $0.25 = \frac{800}{3200}$  . لذلك فإن احتمال الحصول على إطارين معيبين بحسب العلاقة :

$$h(2; 3200, 10, 800) \approx b(2; 10, 0.25)$$

$$= \sum_{x=0}^2 b(x; 10, 0.25) - \sum_{x=0}^1 b(x; 10, 0.25)$$

وباستخدام الجدول II نجد أن :

$$= 0.5256 - 0.244$$

$$= 0.2816$$

يمكن تعميم التوزيع الهندسى الزائدى لمعالجة الحالة التالية :

نفرض أن المجموعة المؤلفة من  $N$  عنصرا تحوى على  $k$  خلية  $A_1, A_2, \dots, A_k$  ، حيث تحوى هذه الخلايا على الترتيب  $a_1, a_2, \dots, a_k$  عنصرا . لنحسب احتمال احتواء عينة حجمها  $n$  مختارة بصورة عشوائية من الـ  $N$  عنصرا على  $x_1$  عنصرا من الخلية  $A_1$  وعلى  $x_2$

عنصرا من الخلية الثانية  $A_2$  وهكذا ... و  $x_k$  عنصرا من الخلية  $A_k$  بحيث يكون :

$$\sum_{i=1}^k x_i = n$$

هذا الاحتمال سنرمز له بالرمز :

$$f(x_1, \dots, x_k; a_1, a_2, \dots, a_k, N, n)$$

الحصول على الصيغة العامة لهذا الاحتمال ، نلاحظ أن عدد الطرق الممكنة لتحقيق هذه التجربة هو  $\binom{N}{n}$  . كما نلاحظ أن هناك  $\binom{a_1}{x_1}$  طريقة لاختيار  $x_1$  عنصرا من عناصر  $A_1$  ، ومن أجل كل طريقة من هذه الطرق هناك  $\binom{a_2}{x_2}$  طريقة لاختيار  $x_2$  عنصرا من عناصر  $A_2$  ، ولذلك فإنه يمكن أن نختار  $x_1$  عنصرا من  $A_1$  و  $x_2$  عنصرا من  $A_2$  بعدد من الطرق مساوٍ لـ  $\binom{a_1}{x_1} \binom{a_2}{x_2}$  طريقة . وبمتابعة هذه العملية نجد أننا نستطيع أن نختار الـ  $n$  عنصرا ، والتي تحوى على  $x_1$  من  $A_1$  و  $x_2$  من  $A_2$  ..  $x_k$  من  $A_k$  بعدد من الطرق مساوٍ لـ  $\binom{a_1}{x_1} \binom{a_2}{x_2} \dots \binom{a_k}{x_k}$  طريقة .

والاحتمال المطلوب يعرف كما يلي :

### تعميم التوزيع الهندسى الزائدى

إذا احتوت مجموعة مؤلفة من  $N$  عنصرا على  $K$  خلية  $A_1, \dots, A_k$  بالعناصر :

$a_1, \dots, a_k$  على الترتيب . يكون توزيع الاحتمال للمتغيرات العشوائية  $X_1, \dots, X_k$  المثلثة لعدد العناصر المختارة من هذه الخلايا بشكل عشوائى من خلال عينة حجمها  $n$  هو :

$$f(x_1, \dots, x_k; a_1, \dots, a_k, N, n) = \frac{\binom{a_1}{x_1} \binom{a_2}{x_2} \dots \binom{a_k}{x_k}}{\binom{N}{n}} \quad \text{حيث :}$$

$$\sum_{i=1}^k a_i = N, \quad \sum_{i=1}^k x_i = n$$

مثال (١٢، ٣) .

سحبنا ، وبشكل عشوائى ، خمسة أوراق من ورق اللعب المؤلف من 52 ورقة . ما هو احتمال أن يكون عدد أوراق الدينارى المسحوبة 2 والبستونى 1 والكبة 2 ؟



## الحل

لنفرض أن :

$X_1 = \{ \text{عدد أوراق الدينارى المسحوبة ضمن الخمس أوراق} \}$

$X_2 = \{ \text{عدد أوراق البستوى المسحوبة ضمن الخمس أوراق} \}$

$X_3 = \{ \text{عدد أوراق الكبة المسحوبة ضمن الخمس أوراق} \}$

$X_4 = \{ \text{عدد أوراق سباني المسحوبة ضمن الخمس أوراق} \}$

نلاحظ أن الورق المؤلف من 52 ورقة يحتوى على أربع خلايا كل خلية تحتوى على 13 ورقة ، هذه الخلايا هي الدينارى ، البستوى ، الكبة والسباني .

والاحتمال المطلوب هو :

$$f(2, 1, 2, 0, 13, 13, 13, 52, 5) = \frac{\binom{13}{2} \binom{13}{1} \binom{13}{2} \binom{13}{0}}{\binom{52}{5}} = 0.03$$

### (٣,٥) التوزيع البواسونى Poisson distribution

للتوزيع البواسونى تطبيقات واسعة . فهو يقدم بصورة خاصة نموذجاً جيداً للمعلومات التي تأخذ شكل التعداد ، حيث يمثل  $X$  فيه عدد الحوادث النادرة ( الملاحظة ) في وحدة معينة زماناً كانت ، أم مسافة ، أم مساحة ، أم حجماً . وكأمثلة على هذا العدد ، نسوق مايلي :

- (١) عدد المكالمات الهاتفية المتبادلة بين الساعة ٤, 5 بين نيويورك وجدة .
- (٢) عدد الذرات الصادرة في ميكروثانية عن كمية من مادة مشعة .
- (٣) عدد الإلكترونات التي يصدرها مهبط مسخن في فترة زمنية محددة .
- (٤) عدد حوادث السير خلال زمن محدد .
- (٥) عدد الأخطاء المطبعية في صفحة ما .
- (٦) عدد باكيتات السجائر المباعة يوم الاثنين في دكان بقالة .
- (٧) عدد البكتريا الموجودة في حجم صغير من سائل معين .
- (٨) عدد ذرات الغاز في منطقة جزئية  $v$  من وعاء حجمه  $V$  .

توضيح الأمثلة السابقة مدى تنوع واتساع تطبيقات التوزيع البواسوني .

إن التجربة التي تقدم لنا قيما عديدة لمثل هذه المتغيرات نطلق عليها اسم تجربة بواسونية ، والملاحظ أن التجربة البواسونية تتمتع بالخواص التالية :

- (١) أن عدد النجاحات التي نحصل عليها في فترة زمنية ( مثلا عدد المكالمات الهاتفية في فترة زمنية ) مستقل عن أى عدد لهذه النجاحات في فترات أخرى .
  - (٢) أن احتمال الحصول على نجاح مفرد في زمن قصير يتناسب مع طول هذا المجال الزمني القصير ولا يعتمد هذا الاحتمال على عدد النجاحات التي نحصل عليها في مجالات زمنية أخرى .
  - (٣) احتمال الحصول على أكثر من نجاح ( في مثل هذا المجال الزمني القصير ) مهمل .
- ملاحظة : يمكن أن تكون الفترة الزمنية المفروضة منطقة معينة ( طول — مساحة — حجم ) .

#### Poisson random variable

#### تعريف (٣,٣) متغير عشوائي بواسوني

إن عدد النجاحات  $X$  ( مثلا عدد المكالمات الهاتفية ) في تجربة يدعى بمتغير عشوائي بواسوني ، كما يدعى توزيع الاحتمال للمتغير  $X$  بالتوزيع البواسوني . سنرمز لهذا النوع بالرمز  $P(x; \mu)$  ، ولقد استخدمنا الرمز السابق عوضا عن  $f(x)$  لكي نشير إلى أن دالة الكثافة الاحتمالية تتعلق بمتوسط عدد النجاحات  $X$  خلال فترة زمنية أو منطقة معينة .

لنفرض في التوزيع الحداني أن  $n$  كبيرة جدا و  $p$  صغيرة جدا ، بحيث إن الجداء  $np = \mu$  يبقى مساويا لعدد ثابت . إذا عوضنا عن  $p$  بـ  $\frac{\mu}{n}$  في عبارة التوزيع الحداني ، فإننا نجد أن :

$$f(x) = \frac{n!}{x! (n-x)!} p^x \cdot q^{n-x}$$

$$= \frac{n!}{x! (n-x)!} \left(\frac{\mu}{n}\right)^x \cdot \left(1 - \frac{\mu}{n}\right)^{n-x}$$

$$\frac{n}{n} \cdot \frac{n-1}{n} \cdot \frac{n-2}{n} \cdots \frac{n-x+1}{n} \cdot \frac{\mu^n (1 - \frac{\mu}{n})^{n-x} (1 - \frac{\mu}{n})^x}{x!}$$

ونلاحظ أن النسب :  $\frac{n-1}{n} \cdot \frac{n-2}{n} \cdots \frac{n-x+1}{n}$  قريبة جدا من الواحد لأن  $n$  كبيرة جدا ، كما فرضنا ، وقيمة  $x$  ثابتة بالنسبة لـ  $n$  . وبأخذ نهاية طرفي العلاقة السابقة نجد أن الطرف الأيمن سيأخذ الشكل التالي :

$$\frac{\mu^x}{x!} \cdot (1 - \frac{\mu}{n})^n$$

ويهرن في الرياضيات أنه عندما تصبح  $n$  كبيرة جدا فإن الكمية  $(1 + \frac{1}{n})^n$  تقترب من عدد نرمز له بالرمز  $e$  ( حيث إن  $e = 2.7183$  ) ، ونعبر عن ذلك بقولنا إن  $e \xrightarrow{n \rightarrow \infty} (1 + \frac{1}{n})^n$  كما ويمكن البرهان على أن  $e^{\mu} \xrightarrow{n \rightarrow \infty} (1 - \frac{\mu}{n})^n$  ، وهكذا نجد أن الشكل الحدى لعبارة  $f(x)$  عندما تسمى  $n$  اللانهاية هو :

$$P(x; \mu) = \frac{\mu^x}{x!} e^{-\mu} , x = 0, 1, 2, \dots, \mu > 0$$

ومن المتوقع أن تعطى هذه العبارة نفس الاحتمالات التي تعطيها عبارة التوزيع الحداني تقريبا ، وذلك شريطة أن تكون  $n$  كبيرة ،  $n.p$  صغيرة نسبيا . هذا ويمكن من أن  $P(x; \mu)$  تحقق شرطى دالة الكثافة الاحتمالية .

### مثال (٣، ١٣)

نتج آلة نوعا معينا من العناصر . إذا علم أن كل 1000 قطعة من إنتاج هذه الآلة تحتوى فى المتوسط على قطعة واحدة معيبة . فاحسب احتمال أن تحتوى عينة من إنتاج هذه الآلة مؤلفة من 8000 قطعة على أقل من 7 قطع معيبة .

### الحل

لنرمز لعدد القطع المعيبة بين الـ 8000 قطعة بالرمز  $x$  ، فيكون المطلوب :

$$\begin{aligned}
 P[X < 7] &= \sum_{x=0}^6 b(x; 8000, 0.001) \\
 &= \sum_{x=0}^6 P(x; 8) \\
 &= 0.3134
 \end{aligned}$$

والملاحظ عند حل هذا المثال أن التجربة الحداثية  $n = 8000$  ،  $P = 0.001$  ، وبما أن  $P$  صغيرة جدا و  $n$  كبيرة . لذلك قربنا التوزيع الحداثي إلى البواسوني باستخدام  $n.p = (8000)(0.001) = 8$

### التوزيع البواسوني

إن توزيع الاحتمال للمتغير العشوائي البواسوني  $X$  والممثل لعدد مرات النجاح التي نحصل عليها خلال فترة زمنية محددة أو منطقة معينة هو :

$$P(x; \mu) = e^{-\mu} \cdot \frac{\mu^x}{x!} : x = 0, 1, 2, \dots$$

حيث يمثل  $\mu$  متوسط عدد مرات النجاح خلال نفس الفترة  $e = 2.7183$  ، يعطينا الجدول III مجموع الاحتمالات البواسونية  $P(x; \mu)$  ، وذلك من أجل بعض قيم من  $\mu = 0.1$  إلى  $\mu = 18$  .

### مثال (٣، ١٤)

يتسلم مقسم هاتف المخابرات الخارجية بين الساعة العاشرة والثانية عشرة بمعدل مخابرتين في الدقيقة .

- (١) لنحسب احتمال ألا يتسلم هذا المقسم أية مخابرة خارجية خلال دقيقة معينة ؟
- (٢) ثم لنحسب احتمال أن يتسلم مخابرتين خلال دقيقة ؟

### الحل

لنرمز لعدد المكالمات الخارجية خلال دقيقة بالرمز  $X$  . نلاحظ أن احتمال ألا يتسلم هذا المقسم أية مخابرة خارجية خلال دقيقة معينة هو :

$$P(X = 0) = e^{-\mu} \cdot \frac{\mu^0}{0!} = e^{-\mu}$$

غير أن معدل المخاطر الخارجية خلال دقيقة  $\mu = 2$  . لذلك فإن :

$$P(X = 0; 2) = e^{-2} = 0.135$$

كما أن احتمال أن يتسلم مخبرتين خلال دقيقة معينة هو :

$$P(X = 2) = P(2; 2) = e^{-2} \cdot \frac{2^2}{2!} = 0.27$$

نظرية (٣, ٤)

إن متوسط وتباين التوزيع البواسوني  $P(x; \mu)$  يساوي  $\mu$  .

البرهان

$$\begin{aligned} E(X) &= \sum_{x=0}^{\infty} x \cdot P(x, \mu) \\ &= \sum_{x=0}^{\infty} x \cdot e^{-\mu} \cdot \frac{\mu^x}{x!} \\ &= \mu \sum_{x=1}^{\infty} \frac{e^{-\mu} \mu^{x-1}}{(x-1)!} \end{aligned}$$

وبفرض أن  $y = x - 1$  نجد :

$$\begin{aligned} &= \mu \cdot \sum_{y=0}^{\infty} e^{-\mu} \cdot \frac{\mu^y}{y!} \\ &= \mu \cdot \sum_{y=0}^{\infty} e^{-\mu} \cdot \frac{\mu^y}{y!} = \sum_{y=0}^{\infty} P(y; \mu) = 1 \end{aligned}$$

لأن :

لنبحث عن :

$$\begin{aligned} E [X (X - 1)] &= \sum_{x=0}^{\infty} x (x - 1) \cdot e^{-\mu} \frac{\mu^x}{x!} \\ &= \mu^2 \cdot \sum_{x=2}^{\infty} \frac{e^{-\mu} \cdot \mu^{x-2}}{(x-2)!} \end{aligned}$$

فإذا فرضنا أن :  $y = x - 2$  فإننا نجد :

$$\begin{aligned} &= \mu^2 \cdot \sum_{y=0}^{\infty} e^{-\mu} \frac{\mu^y}{y!} \\ &= \mu^2 \end{aligned}$$

لذلك فإن :

$$\sigma^2 = E [X (X - 1)] + \mu - \mu^2 = \mu^2 + \mu - \mu^2 = \mu$$

نلاحظ في المثال (٣ و ٤) أن  $\mu = 2$  ، وأن  $\sigma^2 = 2$  أيضا ، لذلك فإن نستطيع باستخدام متباينة تشبيشيف أن نتأكد من أن متغيرنا العشوائى سيقع فى المجال باحتمال على الأقل  $\frac{3}{4}$  . لذلك يمكن القول بأنه خلال ثلاث أرباع الدقيقة سيستقبل المقسم عددا من المكالمات الهاتفية من 0.838 إلى 4.838 مكالمات .

### (٣,٦) التوزيع الحدائى السالب Negative binomial distribution

نفرض أن تجربتنا تتمتع بنفس خواص التجربة الحدائية مع إضافة أننا نكرر التجربة حتى نحصل على عدد معين من النجاحات . لذلك بدلا من البحث عن احتمال الحصول على  $x$  نجاح خلال الـ  $n$  اختبارا ( حيث  $n$  ثابت ) ، فإننا سنبحث عن احتمال الحصول على النجاح ذى الرقم  $k$  فى الاختبار ذى الرقم  $x$  . هذا النوع من التجارب يدعى بالتجارب الحدائية السالبة .

لنحسب مثلاً احتمال الحصول على ثالث صورة في سابع قذفة لقطعة نقود متوازنة ومثالثة . نلاحظ أن هذا الاحتمال هو جداء احتمالين : أولهما احتمال الحصول على صورتين خلال الست قذفات الأولى ، وهذا الاحتمال يساوى :

$$\left(\frac{6}{2}\right) p^2 \cdot q^4$$

والاحتمال الثانى هو احتمال الحصول على وجه الصورة فى القذفة السابقة ويساوى  $p$  .  
وجداء الاحتمالين هو :

$$\left(\frac{6}{2}\right) p^3 \cdot q^4$$

فإذا عوضنا عن  $\frac{1}{2} = p = 1 - q$  ، لوجدنا أن الاحتمال المطلوب يساوى :

$$\left(\frac{6}{2}\right) \left(\frac{1}{2}\right)^3 \cdot \left(\frac{1}{2}\right)^4 = 0.11718$$

### تعريف (٣،٤) المتغير العشوائى الحدافى السالب

#### Negative Binomial Random Variable

إن عدد الاختبارات  $X$  التى تعطى  $k$  نجاحاً فى التجربة الحدافية السالبة يدعى بالمتغير العشوائى الحدافى السالب .

كما أن توزيع الاحتمال للمتغير الحدافى السالب  $X$  يدعى بالتوزيع الحدافى السالب ، وسنرمز له بالرمز  $b^*(x; k, p)$  لأن قيمة هذا التوزيع تتعلق بعدد النجاحات المطلوبة ، وباحتمال النجاح فى كل اختبار . للحصول على الصيغة العامة للاحتمال  $b^*(x; k, p)$  نحسب احتمال الحصول على نجاح فى الاختبار ذى الرقم  $x$  بعد حصولنا على  $k-1$  نجاحاً و  $x-k$  فشلاً فى ترتيب معين . بما أن الاختبارات مستقلة فإنه بإمكاننا ضرب كل الاحتمالات الموافقة لكل نتيجة مطلوبة . وبما أن احتمال الحصول على نجاح فى كل اختبار هو  $p$  وعلى فشل هو  $q = 1-p$  لذلك فاحتمال الحصول على نجاح فى الاختبار ذى الرقم  $x$  وذلك فى الترتيب المعين  $p^k \cdot q^{x-k} = p^k \cdot q^{x-k-1} \cdot q$  ، ويوجد عدد من الترتيبات السابقة يساوى  $\binom{x-1}{k-1}$  ، ولذلك فالاحتمال المطلوب والممثل لاحتمال الحصول على النجاح ذى الرقم  $k$  فى الاختبار ذى الرقم  $x$  يساوى :

$$b^*(x; k, p) = \binom{x-1}{k-1} p^k \cdot q^{x-k} : x = k, k+1, \dots$$

### التوزيع الحدائي السالب

إذا كان احتمال الحصول على نجاح في كل اختبار هو  $p$  ، وعلى فشل هو  $q = 1 - p$  ، فإن توزيع الاحتمال لعدد الاختبارات المجرأة بقصد الحصول على  $k$  نجاح و  $x - k$  فشل يعطى بالعبرة السابقة .

مثال (٣، ١٥)

ما هو احتمال الحصول على صورتين أو كتابتين مرتين عند إلقاء ثلاث قطع نقود سبع مرات ؟

الحل

باستخدام التوزيع الحدائي السالب من أجل  $p = \frac{3}{4}$  ،  $K = 2$  ،  $x = 7$

فإننا نجد أن :

$$b^*(7; 2, \frac{3}{4}) = \binom{6}{1} (\frac{3}{4})^2 (\frac{1}{4})^5 \\ = 0.00109$$

لقد اشتق التوزيع الحدائي السالب اسمه من حقيقة أن كل عنصر في منشور  $(1 - q)^{-k}$  يوافق قيمة من قيم  $b^*(x; k, p)$  من أجل  $x = k, k + 1, k + 2$  لندرس الحالة الخاصة للتوزيع الحدائي السالب من أجل  $k = 1$  ، عندئذ يمثل التوزيع الحدائي السالب توزيع الاحتمال لعدد الاختبارات المجرأة والموافقة لنجاح وحيد . فمثلا في تجربة إلقاء قطعة نقود متوازنة ومتماثلة ، فإننا سنلقى قطعة النقود حتى يظهر وجه الصورة ( نجاح ) لأول مرة ، وعندها نتوقف عن قذف قطعة النقود . سنسحب مثلا احتمال الحصول على وجه صورة في القذفة الرابعة . وفي هذه الحالة سيأخذ التوزيع الحدائي السالب الشكل  $q^{x-1}$  . حيث  $x = 1, 2, 3, \dots$  : س نرمز لهذا التوزيع بالرمز  $g(x; p)$  أى أن احتمال الحصول على صورة في القذفة الرابعة سيكون مساويا لـ :

$$g(4; \frac{1}{2}) = \left(\frac{1}{2}\right) \cdot \left(\frac{1}{2}\right)^{4-1} = \frac{1}{16}$$



## مثال (٣, ١٦)

تحتوى بضاعة مصنعة بالمتوسط بين كل مئة عنصر ثلاثة عناصر معيبة . ما هو احتمال ظهور أول عنصر معيب عند فحص عاشر عنصر من عناصر هذه البضاعة ؟

## الحل

باستخدام التوزيع السابق ، وبفرض أن  $x = 10$ ,  $p = 0.03$ , فإننا نجد :

$$g(10; 0.03) = (0.03) (0.97)^9$$

$$= 0.0228$$

## تمارين محلولة

## تمرين (٩)

ألقينا حجر نرد عشر مرات متتالية . لنعتبر أن النجاح في كل رمية يوافق ظهور أحد العددين 3 أو 5 . ما هو احتمال الحصول على 3 أو 5 خمس مرات ، ثم ما هو توقع وتباين مرات النجاح في هذه التجربة .

## الحل

لنرمز لعدد مرات النجاح بالرمز  $X$  . نجد أن هذا المتغير حداني ، كما نلاحظ أن عدد الاختبارات المجرأة  $n = 10$  ، واحتمال النجاح في كل اختبار هو  $P[3, 5] = \frac{2}{6}$  لذلك فالاحتمال المطلوب هو :

$$P[X = 5] = \binom{10}{5} \left(\frac{1}{3}\right)^5 \left(\frac{2}{3}\right)^5$$

$$= 0.13656$$

أما توقع وتباين عدد مرات النجاح  $X$  فيحسبان بواسطة العلاقتين :

$$\mu = n.p, \sigma^2 = n.p.q$$

أى أن :

$$\mu = 10 \cdot \frac{1}{3} = \frac{10}{3}$$

$$\sigma^2 = \frac{10}{3} \cdot \frac{2}{3} = \frac{20}{9}$$

## تمارين (٢)

أسرة مؤلفة من أب وأم وستة أطفال . إذا علمت أن احتمال أن يكون طفلاً ما في الأسرة ذكراً هو  $\frac{1}{2}$  ، فما هو احتمال أن يكون في الأسرة ثلاثة ذكور ، وثلاث إناث ؟

## الحل

يمكن النظر إلى الأطفال الستة في الأسرة على أساس ست اختبارات ، ويتوافق الذكر نجاح والأنثى فشل . لذلك فإن احتمال أن يكون عدد النجاحات ثلاثة بحسب بالعلاقة :

$$P[X = x] = f(x) = \left(\frac{n}{x}\right) p^x \cdot q^{n-x}$$

وبالتعويض عن  $n = 6$  ،  $p = \frac{1}{2}$  ،  $q = \frac{1}{2}$  ،  $x = 3$  فإننا نجد أن :

$$\begin{aligned} P[X = 3] &= \left(\frac{6}{3}\right) \left(\frac{1}{2}\right)^3 \cdot \left(\frac{1}{2}\right)^3 \\ &= \frac{6!}{3!3!} \cdot \frac{1}{2^6} \\ &= 0.3125 \end{aligned}$$

## تمارين (٣)

يرمى أحمد على هدف بعبارات نارية . إذا علمت أن احتمال إصابته للهدف هو 0.7 فما هو احتمال :

- (١) أن يصيب الهدف مرة واحدة إذا أطلق عليه سبع طلقات متتالية ؟
- (٢) أن يصيب الهدف مرتين على الأقل عندما يطلق سبع طلقات متتالية ؟

## الحل

لنرمز بـ  $X$  لعدد المرات التي أصاب بها الهدف . نجد أن للمتغير  $X$  توزيعاً حدانياً معيناً بالعلاقة :

$$P[X = k] = \binom{7}{k} (0.7)^k \cdot (0.3)^{7-k}$$

(١) أما احتمال إصابته للهدف مرة واحدة عند إطلاق سبع طلقات فهو  
يساوى :

$$P[X = 1] = \binom{7}{1} (0.7) \cdot (0.3)^6 = 0.0035721$$

(٢) لايجاد احتمال أن يصيب الهدف مرتين على الأقل ، علينا أن نبحث عن  
مجموع الاحتمالات من أجل  $k = 2, 3, 4, 5, 6$  ، أو أن نحسب مجموع احتمالات أن  
يصيب مرة وإلا يصيب أى مرة ، ونطرح المجموع من واحد .

نلاحظ أن احتمال عدم الإصابة هو :

$$\begin{aligned} P[X = 0] &= \binom{7}{0} (0.7)^0 (0.3)^7 \\ &= (0.3)^7 \\ &= 0.0002187 \end{aligned}$$

وأخيرا فإن احتمال إصابته مرتين على الأقل هو :

$$\begin{aligned} P[X > 2] &= 1 - P[X < 2] \\ &= 1 - [P(X = 0) + P(X = 1)] \\ &= 1 - 0.0035721 - 0.0002187 \\ &= 0.9962092 \end{aligned}$$

#### تمرين (٤)

أسرة مؤلفة من عشرة أطفال . إذا علمت أن احتمال وجود ذكر يساوى احتمال  
وجود أنثى

- (١) ما هو التوقع الرياضى لعدد الذكور في هذه الأسرة ؟  
(٢) احسب احتمال وجود خمسة ذكور .

## الحل

نلاحظ من الفرض أن  $p = \frac{1}{2}$  ، وأن  $n = 10$  ، وتوزيع الاحتمال لعدد الذكور في الأسرة هو توزيع حداني فيه  $p = \frac{1}{2}$  و  $n = 10$  إذا فالتوقع الرياضي ( لعدد الذكور  $X$  ) هو :

$$E(X) = n.p = 10 \cdot \frac{1}{2} = 5$$

أما احتمال وجود خمسة ذكور فيحسب من العلاقة التالية :

$$\begin{aligned} P[X = 5] &= b(5; 10, \frac{1}{2}) = \binom{10}{5} \left(\frac{1}{2}\right)^5 \left(\frac{1}{2}\right)^5 \\ &= \frac{10!}{5!5!} \cdot \left(\frac{1}{2}\right)^{10} \\ &= 0.24609 \end{aligned}$$

## تمرين (٥)

بفرض أن نسبة اللمبات المعيبة في إنتاج مصنع للأنايب الألكترونية هي 2% ، وإذا علمت أن هذا المصنع قد أنتج 5000 لمبة ، فما هو التوقع الرياضي والانحراف المعياري لعدد اللمبات المعيبة ؟

## الحل

لنرمز بـ  $Y$  لعدد اللمبات المعيبة الموجودة بين 5000 لمبة منتجة . نلاحظ أن للمتغير  $Y$  توزيعا حدانيا باحتمال نجاح قدره  $p = \frac{1}{50}$  ، كما أن  $n = 5000$  لذلك حسب قوانين التوقع والتباين نجد أن :

$$\mu = E(X) = n.p = (5000) \cdot \left(\frac{1}{50}\right) = 100$$

$$\sigma^2 = E(X - \mu)^2 = n.p.q = 100 \cdot \left(\frac{98}{100}\right) = 98$$

والانحراف المعياري كما نعلم ما هو إلا الجذر الموجب للتباين ، ولذلك يكون :

$$\sigma = 9.899$$

### تمرين (٦)

متغير عشوائي حداني  $X$  توقعه 100 وانحرافه المعياري 9.899 . ما هو توزيع هذا المتغير ؟

### الحل

نعلم أن توزيع الاحتمال للمتغير الحداني  $X$  يعطى بالعلاقة :

$$f(x) = \binom{n}{x} p^x \cdot q^{n-x}$$

ولابد من تحديد  $n$  وتحديد  $p$  . لذلك نعلم أنه بالنسبة للمتغير الحداني :

$$\mu = n \cdot p \quad , \quad \sigma = \sqrt{np \cdot (1-p)}$$

وحسب المعطيات نجد أن :

$$100 = n \cdot p \quad , \quad 9.899 = \sqrt{np \cdot (1-p)}$$

وهما معادلتين بمجهولين نجهلها نجد أن  $n = 5000$  ،  $p = \frac{2}{100}$  وبالتعويض في عبارة  $f(x)$  ، فإننا نجد توزيع الاحتمال للمتغير المقروض  $X$  هو :

$$f(x) = \binom{5000}{x} \left(\frac{2}{100}\right)^x \cdot \left(\frac{98}{100}\right)^{5000-x}$$

والملاحظ في هذه المسألة أن عبارة الكثافة الاحتمالية السابقة تمثل توزيع الاحتمال لعدد اللبمبات المعيبة والموجودة ضمن 5000 لمبة إذا علمنا أن نسبة العيب هو  $\frac{2}{100}$  بين إنتاج المصنع من اللبمبات .

## تمرين (٧)

يحتوي صندوق 12 كرة حمراء ، 6 صفراء ، و 8 سوداء . سحبنا وبصورة عشوائية مع الإعادة خمس كرات من الصندوق ، ما هو احتمال أن يكون بين الكرات المسحوبة كرتان حمراء ، كرة صفراء ، وكرتان سوداوات ؟

## الحل

نلاحظ أن الاحتمال المطلوب يمكن إيجاده بواسطة العلاقة :

$$f(x_1, x_2, x_3; p_1, p_2, p_3, n) = \frac{n!}{x_1! x_2! x_3!} p_1^{x_1} \cdot p_2^{x_2} \cdot p_3^{x_3}$$

كما نلاحظ أن :

$$p_1 = \frac{12}{26} \quad p_2 = \frac{6}{26} \quad p_3 = \frac{8}{26} \quad x_1 = 2 \quad x_2 = 1 \quad x_3 = 2 \quad n = 5$$

وبالتعويض نجد أن :

$$f(2, 1, 2; \frac{12}{26}, \frac{6}{26}, \frac{8}{26}, 5) = \frac{5!}{(2!)(1!)(2!)} \left(\frac{6}{13}\right)^2 \cdot \left(\frac{3}{13}\right)^1 \cdot \left(\frac{4}{13}\right)^2$$

$$= 0.13962$$

## تمرين (٨)

قذفنا حجر نرد ست مرات والمطلوب :

(١) ما هو احتمال ظهور الواحد مرة ، الثلاثة مرتين ، والستة ثلاث مرات ؟

(٢) ما هو احتمال ظهور كل وجه مرة واحدة ؟

## الحل

لدينا في هذه المسألة تجربة متعددة الحدود .

(١) الاحتمال المطلوب في (١) يكتب بالشكل التالي :

$$f(1, 2, 3; \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, 6) = \frac{6!}{1!2!3!} \left(\frac{1}{6}\right)^1 \cdot \left(\frac{1}{6}\right)^2 \cdot \left(\frac{1}{6}\right)^3 \\ = 0.001286$$

(٧) أما احتمال ظهور كل وجه مرة واحدة فيسأوى :

$$f(1, 1, 1, 1, 1, 1; \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, 6) = \frac{5}{324} = 0.01543$$

تمثيل (٩)

مجموعة بضاعة مصنعة مؤلفة من مئة عنصر تحتوي على عشرة عناصر معينة .  
سحبنا وبصورة عشوائية عينة مؤلفة من خمس عناصر للدراسة صفتها النوعية . ما هو  
توزيع الاحتمال لعدد العناصر المعنية والموجودة ضمن العينة المستخرجة ؟

الحل

نرمز لعدد العناصر المعنية بـ  $X$  ، نلاحظ أن القيم التي يفترضها المتغير  $X$  هي :

$$0, 1, 2, 3, 4, 5$$

والملاحظ أن هذا المتغير له توزيع هندسي زائدى باحتمال قدره :

$$P[X = K] = \frac{\binom{10}{K} \cdot \left(\frac{90}{5-K}\right)}{\binom{100}{5}} : K = 0, 1, 2, 3, 4, 5$$

وهكذا نجد أن للمتغير  $X$  توزيعاً محدداً بالجدول التالي :

X	0	1	2	3	4	5
P	0.583	0.34	0.07	0.007	0	0

نلاحظ من الجدول السابق أن :

$$\sum_{i=0}^5 P_i = 1$$



## تمرين (١٠)

جملة منتجات مصنعة تخضع للفحص للنظر في إمكانية صلاحيتها للتسويق خارج البلاد ، فإذا علمت أن احتمال اجتياز كل عنصر من هذه المنتجات الفحص ( الاختبار ) هو  $\frac{4}{5}$  ، وأن هذه الاختبارات ( الفحوص ) مستقلة ، وأن عملية الفحص تقف عند ظهور أول عنصر لم يجتز الفحص . فما هو الاحتمال لعدد التجارب المجرأة ؟

## الحل

لنرمز بـ  $Y$  لعدد التجارب المجرأة حتى ظهور أول عنصر لم يجتز الفحص ، نلاحظ أن للمتغير  $Y$  توزيعاً حدانياً سالباً خاصاً وذلك من أجل  $k = 1$  أى أن :

$$P [ Y = y ] = g(y ; p) = \left(\frac{4}{5}\right)^{y-1} \cdot \left(\frac{1}{5}\right) : y = 1, 2, 3, \dots$$

وهكذا نجد أن :

X	1	2	3	.... y ....
P	$\frac{1}{5}$	$\frac{4}{5^2}$	$\frac{(4)^2}{(5)^3}$	$\frac{(4)^{y-1}}{(5)^y}$

## تمرين (١١)

إذا علمت أن احتمال معاناة مريض من رد فعل سيء عند حقنه بمصل معين هو 0.0001 ، فما هو احتمال :

- (١) أن يعاني خمسة بالضبط من رد الفعل السيء من هذا الحقن .
- (٢) أن يعاني أكثر من خمسة من رد الفعل السيء من هذا الحقن .

وذلك عند حقن 3000 مريض بهذا المصل .

## الحل

إذا فرضنا أن  $Y$  يمثل عدد المرضى الذين عانوا من رد الفعل السيء من الحقن

بالموصل المذكور ، عندئذ نلاحظ أن  $Y$  توزيعاً بواسونياً ، ولهذا فلاحتمال يحسب بالعلاقة :

$$P[Y = y] = e^{-\mu} \cdot \frac{\mu^y}{y!}$$

وبالنسبة لاحتمال معاناة خمسة مرضى من رد الفعل السيء فهو يساوى :

$$P[Y = 5] = e^{-\mu} \cdot \frac{\mu^5}{5!} : \mu = n \cdot p = (3000) \left( \frac{1}{10000} \right)$$

ومنه :

$$P[Y = 5] = e^{-\frac{3}{10}} \cdot \frac{\left(\frac{3}{10}\right)^5}{5!} = \frac{1.5}{(10)^5}$$

كذلك فإن احتمال أن يعاني أكثر من خمسة مرضى من رد الفعل السيء هو :

$$P[Y > 5] = 1 - P[Y \leq 5] \\ = 1 - \sum_{y=0}^5 P\left(y; \frac{3}{10}\right)$$

نلاحظ من الجدول III أن :

$$\sum_{y=0}^y P(y; 0.3) = 1$$

إذا

$$\sum_{y=0}^5 p(y, 0.3) = 1$$

أيضا ، والاحتمال المطلوب معلوم .

### تمرين (١٢)

تبلغ نسبة الكؤوس النافقة في إنتاج مصنع زجاج خمسة في المائة من إنتاجه ، ما هو احتمال وجود كأسين تالفين في عينة مؤلفة من 12 كأسا مأخوذة بشكل عشوائى من جملة الإنتاج .

## الحل

نلاحظ أن احتمال ظهور كأس تالف في إنتاج المصنع هو  $P = \frac{5}{100}$  ، فإذا أخذنا عينة من إنتاج المصنع حجمها  $n = 12$  ، وبفرض أن  $X$  يمثل عدد القطع التالفة ضمن هذه العينة ، فإننا نجد أن  $X$  توزيعا بواسونيا نظرا لكبر  $n$  وصغر  $p$  . ولذلك فإن الاحتمال المطلوب هو :

$$\begin{aligned} P[X = 2] &= e^{-\mu} \cdot \frac{\mu^2}{2!} : \mu = n \cdot p = 12 \cdot \frac{5}{100} = 0.6 \\ &= e^{-0.6} \cdot \frac{(0.6)^2}{2!} \\ &= 0.098786 \end{aligned}$$

## تمرين (١٣)

تحتوي مساحة صغيرة من زجاجة مجهرية لفحص الدم من أجل أى شخص طبيعى فى المتوسط على عشرة كريات حمراء . ما هو احتمال أن تحتوى زجاجة من شخص طبيعى فى تلك المساحة الصغيرة على أقل من 6 كريات حمراء ؟

## الحل

لنرمز بـ  $Y$  لعدد الكريات الحمراء التى تحويها المساحة الصغيرة من زجاجة مجهرية لفحص الدم من أجل شخص طبيعى فيكون المطلوب حساب :

$$P[Y < 6] = \sum_{x=0}^5 P(y; \mu)$$

ولكن

$$E(Y) = \mu = 10$$

إذن :

$$P[Y < 6] = \sum_{x=0}^5 P(y; 10)$$

ومن الجدول III نجد أن :

$$= 0.0671$$

### تمارين عامة

(١) أوجد معادلة توزيع الاحتمال للمتغير العشوائى  $X$  الممثل لرقم الكرت المسحوب بشكل عشوائى من صندوق يحوى عشرة كروت مرقعة من الواحد إلى عشرة . ما هو احتمال أن يكون رقم الكرت المسحوب أقل من أربعة ؟

(٢) أوجد التوقع الرياضى والتباين للمتغير العشوائى  $X$  فى التمرين .

(٣) قسمت عجلة الروليت إلى 25 قطاعان بمساحات متساوية من الواحد إلى خمس وعشرين . أوجد عبارة توزيع الاحتمال للمتغير  $X$  الممثل للعدد الذى ظهر عند إدارة عجلة الروليت .

(٤) إذا علمت أن 75% من السيارات التى عبرت نقطة سير ضوئية كانت من داخل مدينة جدة فى المملكة العربية السعودية . ما هو احتمال أن تكون ثلاث على الأقل من ( خمسة سيارات أخرى ستعبر الإشارة الضوئية ) من خارج مدينة جدة ؟

(٥) إذا علمت أن 75% من دجاج مدجنة معينة ملقحة بلقاح ضد مرض يصيب الدجاج . فإذا لقحنا ثلاث دجاجات ، فما هو احتمال أن تكون دجاجتان على الأكثر قد التقطنا المرض ؟

(٦) سحبنا ورقة من ورق اللعب المؤلف من 52 ورقة ثم أعدناها وكررنا التجربة خمس مرات متتالية . ما هو احتمال أن نحصل على ورقتين دينارى وواحدة بستونى ؟

(٧) احسب احتمال الحصول على الأرقام 1, 2, 3, 4, 5, 6 عددا من المرات 1, 2, 3, 1, 2, 4 على الترتيب ، وذلك عند إلقاء حجر نرد عشر مرات متتالية ؟

(٨) سحبنا أربعة صواريخ من مجموعة مؤلفة من عشرة صواريخ وذلك بشكل عشوائى ، وقذفناها بوساطة مدفع . فإذا علمت أن مجموعة العشرة صواريخ تحتوى على ثلاثة صواريخ غير قابلة للانفجار فما هو احتمال :

أ - أن تنفجر الصواريخ الأربعة المختارة ؟

ب - اثنان على الأكثر لم ينفجرا ؟

(٩) في التمرين الثامن . ما هو عدد الصواريخ المعيبة التي تتوقع ألا تنفجر بين الأربعة المسحوبة بشكل عشوائي . استخدم متباينة تشيبيشيف لوصف تغيرات هذا العدد ؟

(١٠) من المقدر أن يكون 4000 صوت من أصوات الناخبين البالغ عددهم 10000 ناخب ضد مشروع فرض ضريبة على الأرباح . فإذا اخترنا 15 شخصا ممن يحق لهم التصويت وتم سؤالهم عن رأيهم ، فما هو احتمال أن يكون سبعة على الأكثر مع مشروع فرض الضريبة المذكورة ؟

(١١) إذا علمت أن متوسط عدد حوادث السير خلال أسبوع عند إشارة ضوئية معينة هو ثلاثة ، فما هو احتمال أن تقع خمسة حوادث بالضبط عند هذه الإشارة خلال الأسبوع القادم ؟

(١٢) تم تشكيل اتحاد للطلاب المسلمين من أربعة أقطار عربية ، وقد انضم إلى هذا الاتحاد ثلاثة طلاب سعوديين ، خمسة فلسطينيين ، وطلابين سوريين وطلابين مصريين . ما هو احتمال أن تمثل الجنسيات الأربعة في لجنة مؤلفة من أربعة طلاب مختارة بشكل عشوائي ؟

(١٣) إذا علمت أن متوسط عدد الأخطاء التي ترتكبها سكرتيرة عند طباعتها الورقة على الآلة الكاتبة هو اثنان ، فما هو احتمال :

أ - أن ترتكب هذه السكرتيرة أربعة أخطاء أو أكثر في الصفحة الثانية .

ب - ألا ترتكب أى خطأ في الصفحة الثانية .

(١٤) إذا علمت أن احتمال موت شخص في عدوى تصيب الجهاز التنفسي هو 0.002 . أوجد احتمال أن يكون أكثر من خمسة أشخاص من أصل 2000 شخص مصابين بهذه العدوى قد ماتوا .

(١٥) ما هو احتمال الحصول على صورة نائلة في الالقاء السابعة لدى إلقاء قطعة نقود سبع مرات متتالية ؟

(١٦) للحصول على رخصة قيادة يتقدم المواطن للفحص أمام لجنة ، فقد يجتاز

هذا الفحص وقد يفشل . إذا علمت أن احتمال أن اجتياز مواطن الفحص في كل مرة هو 0.7 ، فما هو احتمال أن يجتاز سعيد فحص القيادة هذا في :

أ - المرة الثالثة .

ب - قبل المرة الرابعة .

## الفصل الرابع

### بعض توزيعات الاحتمال المستمرة

■ التوزيع الطبيعي ■ المساحة تحت المنحنى الطبيعي ■ التقريب الطبيعي  
للتوزيع الحثائي ■ التوزيعات غاما ، الأسي ، كاي-مربع ■ توزيع  
وايل ■ تقارين محلولة ■ تقارين عامة .





## (٤,١) التوزيع الطبيعي Normal distribution

### تمهيد

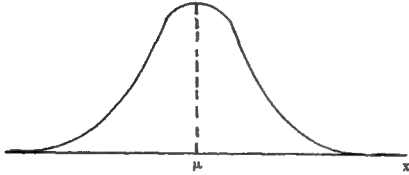
كما ذكرنا سابقا فإن المتغير العشوائى المستمر يسمح بجميع نقاط محور محوره ، وبالتالي فإنه بالإضافة لكون عدد قيم هذا المتغير لانهائياً ، فإن هذا العدد غير قابل للعد كنقاط المجال (a, b) مثلاً . وكبعض الأمثلة على المتغيرات المستمرة نذكر على سبيل المثال عمر مصباح كهربائى ، أخطاء القياسات فى تجربة بحرية ، طول إنسان ... إلخ للحصول على نموذج احتمالى لمتغير مستمر نبدأ باختبار منحني مستمر يمثل ما يسمى بدالة الكثافة الاحتمالية  $f(x)$  . ولا بد لهذه الدالة أن تحقق الشرطين التاليين :

$$1. f(x) \geq 0 \text{ من أجل جميع قيم } x$$

٢) المساحة بين المنحنى  $f(x)$  والمحور السينى تساوى الواحد .

وعندئذ سيكون احتمال أى حادث ممثلاً للمساحة تحت منحنى الكثافة . ونتيجة لذلك ، فإن احتمال فرض أن المتغير  $X$  مساوياً للقيمة  $a$  مثلاً  $P(X = a)$  هى المساحة تحت المنحنى فوق النقطة  $a$  من المحور السينى وهذه المساحة صفر . وهكذا نجد أن احتمال أن يأخذ المتغير العشوائى المستمر ( قيمة بحد ذاتها ) معلوم . وهذا تعبير واقعى عن استحالة توصيل الإنسان إلى أجهزة قياس دقيقة بصورة مطلقة . لذلك تبقى مثل هذه النتيجة مقبولة طالما بقى الإنسان غير قادر على الادعاء بأن قياساته لعمر مصباح كهربائى ، مثلاً ، هى قياسات لا تخضع لأى خطأ على الإطلاق ، إذ مهما أوتى جهاز القياس من الدقة ، ومهما بلغت مهارة الإنسان الذى يستخدم هذا الجهاز فلا بد من ارتكاب خطأ مهما كان صغيراً .

وبينما نتخذ منحنيات الكثافة أشكالاً مختلفة ، نلاحظ أن عدداً كبيراً من المتغيرات العشوائية التى تصادفنا فى تجاربنا اليومية لها منحنى كثافة له تقريباً شكل الجرس ، أو كما نعبّر عن ذلك إحصائياً ، له بصورة تقريبية ، شكل التوزيع الطبيعى كما هو موضح فى الشكل (٤,١) .



الشكل (٤,١)

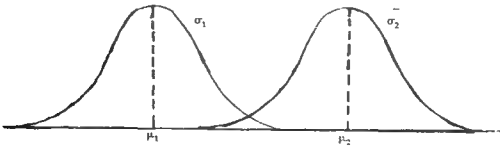
يدعى المتغير العشوائى ذو التوزيع الجرسى كما فى الشكل (٤,١) بالمتغير الطبيعى ، والمعادلة الرياضية لهذا التوزيع تتعلق بمحولين  $\sigma, \mu$  ( أى المتوسط والانحراف المعيارى للمتغير الطبيعى ، لذلك سنرمز للكثافة الاحتمالية لهذا المتغير بالرمز  $n(x; \mu, \sigma)$  )

### التوزيع الطبيعى

إن الكثافة الاحتمالية للمتغير العشوائى الطبيعى  $X$  ذى التوقع  $\mu$  ، والانحراف المعيارى  $\sigma$  تعطى بالعلاقة :

$$n(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi} \cdot \sigma} e^{-\frac{1}{2} \left( \frac{x-\mu}{\sigma} \right)^2} ; -\infty < x < +\infty$$

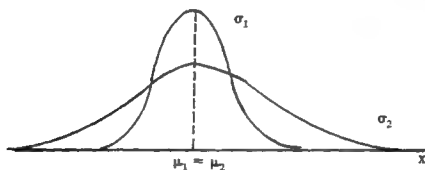
حيث يمثل  $e = 2.71828 = 3.14159$  . يكون المنحنى الطبيعى محدداً تماماً إذا كانت  $\sigma, \mu$  محددين . مثلاً إذا كان  $\mu = 36, \sigma = 2$  عندئذ يمكن بسهولة حساب  $n(x; 36, 2)$  لجميع قيم  $x$  ويمكن رسم هذا المنحنى . يوضح الشكل (٤,٢) منحنين طبيعيين لهما نفس الانحراف المعيارى  $\sigma$  بتوقعين مختلفين .



الشكل (٤,٢)

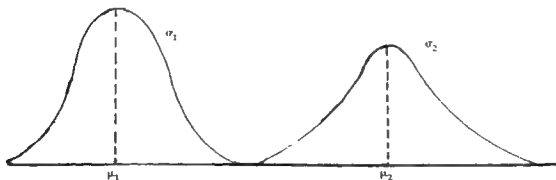
ونلاحظ أن المنحنيين متطابقان بالشكل ، ولكنهما متمركزين في نقطتين مختلفتين على طول محور الفواصل .

يوضح الشكل (٤,٣) منحنيين طبيعيين لهما نفس التوقع (الوسط) ، وانحرافين مختلفين ، كما نلاحظ على نفس الشكل أن المنحنيين متمركزان في نفس نقطة على محور الفواصل ، ولكن المنحنى ذو الانحراف المعيارى الأكبر أخفض من المنحنى الآخر .



الشكل (٤,٣)

أما الشكل (٤,٤) فيوضح الرسم البياني لمنحنيين طبيعيين لهما توقعين مختلفين وانحرافين مختلفين أيضا ، ومن الواضح أنهما متمركزان في نقطتين مختلفتين على محور الفواصل وشكليهما يعكسان اختلاف انحرافيهما المعيارين .



الشكل (٤,٤)

$$\mu_1 > \mu_2 , \sigma_1 < \sigma_2$$

بالعودة إلى المشتقتين الأولى والثانية للدالة  $n(x; \mu, \sigma)$  مقارنة الأشكال السابقة نستنتج الخواص التالية التي يتمتع بها المنحنى الطبيعي .

- (١) يبلغ المنحنى نهايته العظمى في النقطة  $x = \mu$  من المحور السيني .
  - (٢) المنحنى الطبيعي متناظر حول محور الترتيب المار من النقطة  $\mu$  .
  - (٣) يعاني المنحنى الطبيعي انعطافا في النقاط  $x = \mu \pm \sigma$  ، كما أن تقعره يكون نحو الأسفل إذا كان  $\mu - \sigma < X < \mu + \sigma$  . ونحو الأعلى فيما عدا ذلك .
  - (٤) يصبح المنحنى الطبيعي مقاربا من محور الفواصل كلما ابتعدنا عن  $\mu$  في مختلف الاتجاهات .
  - (٥) المساحة تحت المنحنى الطبيعي وفوق محور الفواصل تساوى الواحد .
- سنبرهن أن  $\mu$  ,  $\sigma^2$  يمثلان على الترتيب توقع وتباين المتغير العشوائى الطبيعي .  
ولحساب التوقع الرياضى نكتب الآتي :

$$E(X) = \frac{1}{\sqrt{2\pi} \cdot \sigma} \int_{-\infty}^{+\infty} x \cdot e^{-\frac{1}{2} \left( \frac{x-\mu}{\sigma} \right)^2} dx$$

وباجراء تغيير للمتحول بالعلاقة  $Z = \frac{x-\mu}{\sigma}$  نجد أن  $dx = \sigma \cdot dz$  و  $X = \sigma \cdot Z + \mu$  وبالتعويض نجد أن :

$$\begin{aligned} E(X) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} (\sigma \cdot Z + \mu) \cdot e^{-\frac{1}{2} Z^2} dz \\ &= \frac{\mu}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{1}{2} Z^2} dz + \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} Z e^{-\frac{1}{2} Z^2} dz \end{aligned}$$

ومن الواضح أن التكامل  $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{1}{2} Z^2} dz$  يمثل المساحة الواقعة تحت المنحنى الطبيعي ذى التوقع صفر والانحراف المعياري واحد لذلك فقيمته الواحد . ومنه  $\mu \cdot 1 = \mu$  ، أما التكامل الثانى فيبرهن أنه يساوى الصفر . وهكذا نجد أن :

$$E(X) = \mu \cdot 1 + 0 = \mu$$

أما بالنسبة للتباين ، فمن المعلوم أن :

$$E[(X - \mu)^2] = \frac{1}{\sqrt{2\pi} \cdot \sigma} \int_{-\infty}^{+\infty} (x - \mu)^2 \cdot e^{-\frac{1}{2} \left( \frac{x - \mu}{\sigma} \right)^2} dx$$

وبتغيير المتحول بوساطة العلاقة السابقة نفسها نجد أن :

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \sigma^2 \cdot Z^2 \cdot e^{-\frac{1}{2} Z^2} dz$$

$$= \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} Z^2 \cdot e^{-\frac{1}{2} Z^2} dz$$

وبالتكامل بالتجزئة على اعتبار أن  $U = Z$  و  $dv = Z \cdot e^{-\frac{1}{2} Z^2} dZ$  نجد أن :

$$E[(X - \mu)^2] = \frac{\sigma^2}{\sqrt{2\pi}} \left[ -Z \cdot e^{-\frac{Z^2}{2}} \Big|_{-\infty}^{+\infty} + \int_{-\infty}^{+\infty} e^{-\frac{1}{2} Z^2} dz \right]$$

$$= \frac{\sigma^2}{\sqrt{2\pi}} [0 + \sqrt{2\pi}]$$

$$= \sigma^2$$

## (٤, ٢) المساحة تحت المنحنى الطبيعي

### Area under the normal distribution

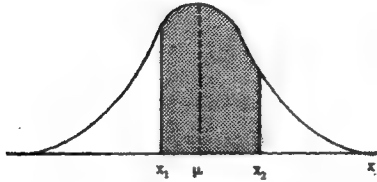
من المعلوم أن المساحة تحت منحنى التوزيع الاحتمالي المستمر ، وبين أى نقطتين  $x = x_2$   $x = x_1$  يمثل احتمال أن يتردد المتغير المفروض بين هاتين النقطتين أى أن :

$$P[x_1 < X < x_2] = \int_{x_1}^{x_2} f(x) dx$$

ومن أجل التوزيع الطبيعي نجد أن :

$$P[x_1 < X < x_2] = \frac{1}{\sqrt{2\pi} \cdot \sigma} \int_{x_1}^{x_2} e^{-\frac{1}{2} \left( \frac{x - \mu}{\sigma} \right)^2} \cdot dx$$

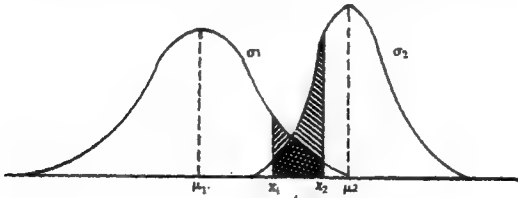
وهذا التكامل يمثل مساحة القسم المظلل من الشكل (٤, ٥) .



$$P[x_1 < X < x_2] = \text{مساحة القسم المظلل}$$

الشكل (٤,٥)

لقد رأينا من خلال الأشكال (٤,٢) ، (٤,٣) ، (٤,٤) كيف أن المنحنى الطبيعي يعتمد على التوقع والانحراف المعياري للتوزيع ، وأن المساحة تحت المنحنى بين أي نقطتين تعتمد أيضا قيم كل من  $\sigma$  ،  $\mu$  ، وهذا ما يوضحه الشكل (٤,٦) ، حيث ظللنا المنطقة المقابلة لـ  $P[x_1 < X < x_2]$  حيث يمثل  $X$  متغيراً عشوائياً يصفه التوزيع I يمثل القسم المظلل من نفس الشكل . أما إذا كان  $X$  متغيراً عشوائياً موصوفاً بالتوزيع II ، فإن هذا الاحتمال يمثل القسم المظلل من نفس الشكل II . وبما أن المساحتين المظلتين على كل شكل مختلفتان ، لذلك فإن الاحتمال سيختلف تبعاً لشكل توزيع المتغير  $X$  .



الشكل (٤,٦)

ولحسن الحظ أنه يمكننا أن نغير كافة الملاحظات المتعلقة بمتغير عشوائي ما  $X$  إلى مجموعة ملاحظات متغير عشوائي آخر  $Z$  ذو التوقع صفر والتباين واحد . ويمكن إجراء ذلك بواسطة علاقة التحويل :

$$Z = \frac{X - \mu}{\sigma}$$

فإذا افترض المتغير  $X$  القيمة  $x$  ، فنندئذ يفترض المتغير  $Z$  القيمة  $\frac{x - \mu}{\sigma}$  .  
 إذا وقع  $X$  بين القيمتين  $x_1, x_2$  فإن المتغير الجديد  $Z$  سيقع بين القيمتين  $\frac{x_1 - \mu}{\sigma}$  و  $\frac{x_2 - \mu}{\sigma}$  على الترتيب . ولهذا يمكن أن نكتب :

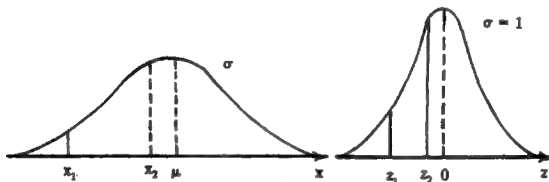
$$\begin{aligned} P [x_1 < X < x_2] &= \frac{1}{\sqrt{2\pi}\sigma} \int_{x_1}^{x_2} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \cdot dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{\frac{x_1-\mu}{\sigma}}^{\frac{x_2-\mu}{\sigma}} e^{-\frac{1}{2}z^2} dz \\ &= \int_{\frac{x_1-\mu}{\sigma}}^{\frac{x_2-\mu}{\sigma}} n(z; 0, 1) dz = P [z_1 < Z < z_2] \end{aligned}$$

حيث رمزنا بـ  $z$  للمتغير الطبيعي ذى التوقع الصفري والتباين المساوى للواحد .

#### تعريف (٤,١) التوزيع الاحتمالى Probability distribution

يدعى التوزيع الاحتمالى للمتغير العشوائى الطبيعي بالتوقع صفر والتباين واحد بالتوزيع الطبيعي المعيارى .

يوضح الشكل (٤,٧) كلا من التوزيعين الطبيعي والطبيعى المعيارى .



الشكل (٤,٧)

يعطينا الجدول IV المساحة الواقعة أسفل التوزيع الطبيعي المعيارى الموافقة لاحتمال  $P[Z < 2]$  . وذلك من أجل جميع قيم  $Z$  من 3.4 - إلى 3.4 . سنذكر فيما يلى طرق استخدام هذا الجدول .

## مثال (٤,١)

$X$  متغير عشوائى طبيعى بالتوقع  $\mu = 36$  ، والتباين  $\sigma^2 = 4$  ، لحساب  $P[30 < X < 37]$  ، أى المساحة الواقعة تحت المنحنى الطبيعى المعيارى السابق بين النقطتين 30, 37 ، فإننا نقوم بتحويل هذا التوقع إلى التوزيع الطبيعى بواسطة علاقة التحويل :

$$Z = \frac{x - \mu}{\sigma} = \frac{x - 36}{2}$$

نلاحظ أنه يقابل النقطتين  $x_1 = 30$  ،  $x_2 = 37$  النقطتان :

$$Z_1 = \frac{30 - 36}{2} = -3$$

$$Z_2 = \frac{37 - 36}{2} = 0.5$$

لذلك فإن :

$$\begin{aligned} P[30 < X < 37] &= P[-3 < Z < 0.5] \\ &= P[Z < 0.5] - P[Z < -3] \end{aligned}$$

وباستخدام الجدول IV من أجل  $Z = 0.5$  ، من أجل  $-3$  نجد أن :

$$\begin{aligned} &= 0.3085 - 0.0013 \\ &= 0.2072 \end{aligned}$$

## ملاحظة هامة

بفرض أن للمتغير العشوائى الطبيعى  $X$  توقعاً  $\mu$  وتبايناً  $\sigma^2$  فإننا نلاحظ أن :



$$\begin{aligned}
 P[\mu - 2\sigma < X < \mu + 2\sigma] &= P\left[\frac{\mu - 2\sigma - \mu}{\sigma} < \frac{X - \mu}{\sigma} < \frac{\mu + 2\sigma - \mu}{\sigma}\right] \\
 &= P[-2 < Z < 2] \\
 &= P[Z < 2] - P[Z < -2]
 \end{aligned}$$

ومن الجدول IV ، نجد أن : من أجل  $Z = 2$  و  $Z = -2$  :

$$P[\mu - 2\sigma < X < \mu + 2\sigma] = 0.9772 - 0.0228$$

$$P[\mu - 2\sigma < X < \mu + 2\sigma] = 0.9544$$

والمتباينة الأخيرة أفضل وأقوى من متباينة تشبيشيف . نخلص للقول بأن شكل متباينة تشبيشيف من أجل المتغيرات الطبيعية بالتوقع  $\mu$  والانحراف المعياري  $\sigma$  تأخذ الشكل الجديد التالي :

$$P[\mu - 2\sigma < X < \mu + 2\sigma] = 0.9544$$

### مثال (٤،٢)

رُكِبَتْ في إحدى القواعد الحربية مدفعية أرضية ، فإذا علمت أن عمر هذه المدفعية ( أى صلاحيتها للعمل ) يتوزع وفقاً للتوزيع الطبيعي بالمتوسط 4 سنوات ، والانحراف المعياري 0.6 سنة ، فاحسب احتمال أن تظل هذه المدفعية صالحة للعمل في أقل من 3.3 سنوات .

### الحل

لنرمز لفترة عمل المدفعية السابقة بالرمز  $X$  . نلاحظ أن  $Z = \frac{X - 4}{0.6}$  يمثل متغيراً طبيعياً معيارياً والاحتمال المطلوب هو :

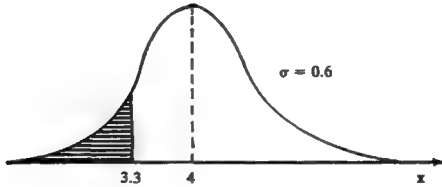
$$P[X < 3.3] = P\left[Z < \frac{3.3 - 4}{0.6}\right]$$

$$= P[Z < -1.17]$$

وبالعودة إلى الجدول IV ، نجد أنه يقابل القيمة  $Z = -1.17$  مساحة قدرها 0.1210 تحت المنحنى الطبيعي المعياري ، ولذلك فلاحتمال المطلوب هو :

$$P[X < 3.3] = 0.1210$$

يوضح الشكل (٤،٨) هذه المساحة بالقسم المظلل من الشكل .



الشكل (٤،٨)

مثال (٤،٣)

بفرض أن لنوع معين من المصابيح الكهربائية عمراً يتوزع بشكل طبيعي بالوسط 750 ساعة ، والانحراف المعياري 60 ساعة . ما هو احتمال احتراق لمبة من هذه اللمبات بين 700 و 780 ساعة من اشتغالها ؟

الحل

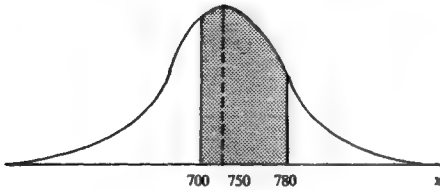
يوضح الشكل (٤،٨) توزيع عمر هذا النوع من المصابيح المنتجة ، ونلاحظ أن قيم  $Z$  الموافقة لـ  $x_1 = 700$  ،  $x_2 = 780$  هي :

$$Z_1 = \frac{700 - 750}{60} = -0.83$$

$$Z_2 = \frac{780 - 750}{60} = 0.5$$

لذلك فالاحتمال المطلوب هو :

$$\begin{aligned}
 P[700 < X < 780] &= P[-0.83 < Z < 0.5] \\
 &= P[Z < 0.5] - P[Z < -0.83] \\
 &= 0.6915 - 0.2033 \\
 &= 0.4882
 \end{aligned}$$



الشكل (٤,٩)

مثال (٤,٤)

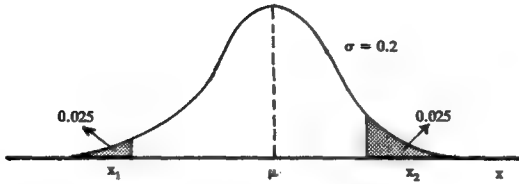
استخدمت مجموعة من القياسات لرفض جميع المركبات غير الموجودة في المجال  $(1.5 - d, 1.5 + d)$  لبعد معين . فإذا عُلِمَ أن هذه القياسات توزيعاً طبيعياً بالمتوسط 1.5 والانحراف المعياري 0.2 . فما هي قيمة  $d$  بحيث يغطي المجال المذكور 95% من هذه القياسات ؟

الحل

لنحدد قبل كل شيء قيمة كل من  $Z_2, Z_1$  بحيث يكون :

$$P(Z_1 < Z < Z_2) = 0.95$$

نلاحظ من الشكل (٤,١١) ومن الجدول IV أن  $Z_2 = 1.96, Z_1 = -1.96$



الشكل (٤,١٠)

لنحدد الآن قيمة  $x_2 = \sigma \cdot Z + \mu$  بحيث تكون :

$$x_2 = 1.5 + d = (0.2) (1.96) + 1.5$$

من هذه العلاقة نجد أن  $d = 0.392$

#### مثال (٤,٥)

تنتج آلة معينة مقاومات كهربائية لها وسط مقاومة 35 أوم وانحراف معيارى قدره 1.8 أوم . بفرض أن المقاومة تتوزع طبيعياً ، ويمكن أن تقاس بالنسبة لأى درجة من الدقة ، فما هى النسبة المئوية للمقاومات التى سيكون لها مقاومة أكبر من 38 أوم ؟

#### الحل

نلاحظ أنه يمكن إيجاد النسبة المئوية بضرب التكرار النسبى بـ 100% . وبما أن التكرار النسبى من أجل أى مجال يساوى احتمال الوقوع فى هذا المجال ، لذلك يجب إيجاد المساحة الواقعة على يمين النقطة  $x = 38$  كما هو موضح على الشكل (٤,١١) وهذا بالطبع

$$Z = \frac{38 - 35}{1.8} \text{ القيمة الموافقة } x = 38 \text{ إلى القيمة الموافقة}$$

$$P[X > 38] = P[Z > 1.67]$$

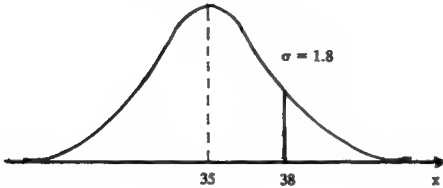
$$= 1 - P[Z < 1.67]$$

ومن الجدول IV نجد أن :

$$= 1 - 0.9527$$

$$= 0.0473$$

لذلك فإن 4.7% من المقاومات سيكون لها مقاومة تفوق 38 أوم .



الشكل (٤,١١)

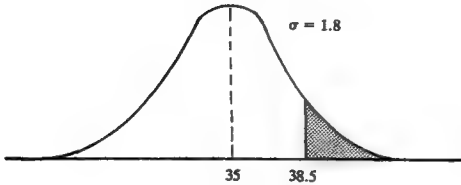
مثال (٤,٦)

ما هي النسبة المئوية للمقاومات التي تفوق 38 أوم في المثال السابق إذا علمت أن المقاومة مقاسة بالنسبة لأقرب أوم ؟

الحل

جوهر هذا المثال يختلف عن المسألة الواردة في المثال (٤,٥) ففي المثال السابق ربطنا بكل مقاومة قياساً قدره 38 أوم . بالنسبة لكل المقاومات التي لها قيمة أكبر من 37.8 وأقل 38.5 أوم . ونحن متأكدون من تقريب التوزيع المنقطع باعتباره توزيعاً طبيعياً مستمراً . والمساحة المطلوبة هي المنطقة المظللة على يمين النقطة 38.5 من الشكل (٤,١٢) . ونجد أن :

$$Z = \frac{38.5 - 35}{1.8} = 1.944$$



الشكل (٤, ١٢)

لذلك فإن :

$$\begin{aligned}
 P[X > 38.5] &= P[Z > 1.944] \\
 &= 1 - P[Z < 1.944] \\
 &= 1 - 0.9738 \\
 &= 0.262
 \end{aligned}$$

من هذا نستنتج أن 2.6% من المقاومات تفوق 38 أوم عندما تقاس بالنسبة لأقرب أوم .  
وعمل الفرق  $4.7\% - 2.6\% = 2.1\%$  بين الإجابتين في المثالين كل المقاومات التي لها قيم أكبر من 38 وأقل من 38.5 أوم والتي اعتبرناها تساوى 38 أوم .

### (٤, ٣) التقريب الطبيعي للتوزيع الحداني

#### Normal Approximation to the Binomial

رأينا في الفصل السابق تطبيقات متعددة للتوزيع الحداني ، وحسبنا فيها جميعا احتمال أن يأخذ المتغير  $X$  ( الممثل لعدد النجاحات في  $n$  اختباراً ) قيمة معينة ، وقد اقتصرنا في دراستنا على قيم  $n$  الصغيرة ، بسبب المشتقة في حساب  $b(x; n, p)$  عندما تكون  $n$  كبيرة ( وفي مثل هذه الحالة علينا أن نحسب الاحتمال الحداني بواسطة إجراءات التقريب ) ودرسنا أيضا في الفقرة (٣, ٥) طريقة تقريب التوزيع الحداني إلى التوزيع الطبيعي عندما تكون  $n$  كبيرة و  $p$  قريبة من الصفر أو الواحد ، ولاحظنا عندئذ أن كلا

من التوزيع البواسوني والحداني هو توزيع منقطع ، وفي المثال (٤,٦) كنا قد بينا أول تطبيق حول تقريب التوزيع الحداني إلى الطبيعي . سنذكر فيما يلي نظرية تدعى بنظرية النهايات المركزية تقرب لنا التوزيع الحداني إلى التوزيع الطبيعي . وذلك في الحالة التي تكون فيها  $n$  كبيرة بقدر كاف و  $p$  ليست قريبة من الصفر أو الواحد ، وسنذكر هذه النظرية بدون برهان .

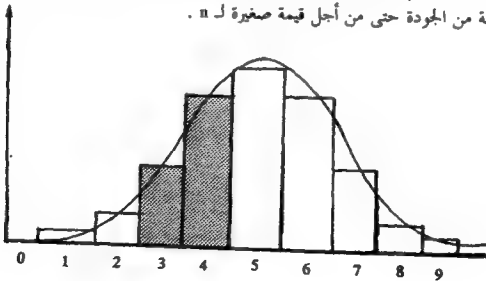
### نظرية (٤,٩)

إذا كان  $X$  متغيراً عشوائياً حدانياً بالتوقع  $\mu = np$  ، والتباين  $\sigma^2 = npq$  عندئذ يكون للمتغير :

$$Z = \frac{X - np}{\sqrt{np.q}}$$

توزيع يسعى إلى التوزيع الطبيعي المعياري وذلك عندما تسعى  $n$  إلى  $\infty$  .

على سبيل المثال ، لنعتبر التوزيع الحداني من أجل  $n = 10$  و  $p = \frac{1}{2}$  . عندئذ نجد أن  $\mu = np = 5$  و  $\sigma = \sqrt{np.q} = 1.58$  . يبين الشكل (٤,١٣) الاحتمال الموافق لحادث معين ، وذلك عند استخدام كل من التوزيع الحداني والتوزيع الطبيعي ، كما تحدده نظرية النهايات المركزية . ونظرة أولية للشكل يبين أن التقريب جيد تماماً حتى في الحالة  $n = 10$  . ويسهم تناظر التوزيع الحداني في حالة  $n = 10$  في كون التقريب على هذه الدرجة من الجودة حتى من أجل قيمة صغيرة لـ  $n$  .



الشكل (٤,١٣)

واحتيال أن يكون  $X$  مساوياً لـ 3 أو 4 يساوى تماماً مساحة المستطيلين المقامين فوق  $x = 3$  و  $x = 4$  . ويمكن تقريب هذه المساحة تحت المنحنى الطبيعي من  $x = 2.5$  إلى  $x = 4.5$  وهي المساحة المظللة في الشكل (٤،١٣) .

والملاحظ أنه عندما تكون  $n$  صغيرة و  $p$  قريبة من الصفر أو الواحد ، فإن شكل المضلع الاحتمالي سيكون منحازاً ( أى يتجمع معظمه ) إلى جانب القيمة  $x = 0$  أو  $x = n$  على الترتيب . أى إنه سيكون بعيداً جداً عن وضع التناظر . وفي مثل هذه الحالات سيكون التقريب سيئاً . وبصورة عامة كلما ابتعدت  $p$  عن القيمة  $\frac{1}{2}$  كلما ابتعد شكل المضلع الاحتمالي للتوزيع الحدى عن التناظر .

وإذا تذكرنا ما أشرنا إليه من أن 95% تقريباً من قياسات المجتمع الطبيعي ستقع ضمن المجال  $(\mu \pm 2\sigma)$  ، فإننا نستنتج أن القياسات في المجتمع الحداني ستكون منتشرة فوق المجال  $(\mu \pm 2\sigma)$  . ولتحديد متى سيكون التقريب باستخدام التوزيع الطبيعي مناسباً أم لا ، نحسب  $\mu = np$  و  $\sigma = \sqrt{npq}$  فإذا وقع المجال  $\mu \pm 2\sigma$  ضمن مدى التوزيع الحداني ، أى بين 0 و  $n$  . فسيكون التقريب جيداً .

#### مثال (٤،٧)

لنحسب في التجربة الحدانية الموضحة على الشكل (٤،١٣) حيث فرضنا  $n = 10$  ،  $p = \frac{1}{2}$  احتمال أن يكون  $X$  مساوياً 2 أو 3 أو 4 بدقة تصل الرقم العشري الرابع ، وذلك باستخدام جدول التوزيع الحداني II ، ثم باستخدام التقريب الطبيعي للتوزيع الثنائي .

#### الحل

نلاحظ أن :

$$P_1 = P[X = 2 \text{ أو } X = 3 \text{ أو } X = 4] = \sum_{x=0}^4 b\left(x; 10, \frac{1}{2}\right) - \sum_{x=0}^1 b\left(x; 10, \frac{1}{2}\right)$$

$$= 0.3770 - 0.0108$$



وهكذا نجد أن :

$$P_1 = 0.3662$$

أما إذا استخدمنا التقريب الطبيعي للتوزيع الحداني ، فإننا نقوم بحساب المساحة بين  $x_1 = 1.5$  ،  $x_2 = 4.5$  . وهكذا نكتب :

$$P_2 = P[1.5 < X < 4.5] = P \left[ \frac{1.5 - 5}{1.58} < \frac{X - \mu}{\sigma} < \frac{4.5 - 5}{1.58} \right]$$

$$= P[-2.22 < Z < 0.32]$$

$$= P[Z < -0.32] - p[Z < -2.22]$$

$$= 0.4868 - 0.1255 = 0.3613$$

$$P_2 = 0.3613$$

وبذلك نحصل على :

ونلاحظ أن القيمة  $P_2$  تتفق مع القيمة الحقيقية  $P_1$  برقمين عشريين .

### مثال (٤,٨)

اخترنا لقاحاً ضد الزكام . وقد أعطى اللقاح لمتى شخص ، وتمت مراقبتهم بالنسبة لإصابتهم بالزكام لمدة عام ، وقد نجا منهم 120 شخصاً من الإصابة . فإذا فرضنا أن احتمال عدم الإصابة بالزكام بصورة طبيعية دون استخدام أى لقاح هو 0.5 ، فما هي النتيجة التي يمكن استخلاصها من هذه التجربة حول فعالية اللقاح ؟

### الحل

لنرمز بـ  $p = 0.5$  ،  $n = 200$  . باستخدام التقريب الطبيعي للتوزيع الحداني لنرمز بـ  $X$  لعدد الناجين من الإصابة بالزكام ، ولنحسب احتمال أن يكون  $X$  أكبر من أو يساوي

نلاحظ أن :

$$\sigma = \sqrt{np.q} = \sqrt{(200) (0.5) (0.5)} = 7.07$$

$$\mu = np = 200 (0.5) = 100$$

$$P [X \geq 120] = P \left[ \frac{X - \mu}{\sigma} \geq \frac{120 - 100}{7.07} \right]$$

$$= P [Z \geq 2.828]$$

$$= 1 - P [Z < 2.828]$$

$$= 1 - 0.9977$$

كما نلاحظ أيضا هذا الاحتمال من الصفر بحيث يمكن إهماله ، وهذا يدعونا إلى الاعتقاد بأن اللقاح مفيد بالفعل في منع الإصابة بالزكام .

#### مثال (٤,٩)

تحتوي مجموعة بضاعة مصنعة على خمسة بالمئة من القطع المعيبة . فإذا اخترنا وبصورة عشوائية مئة قطعة من هذه المجموعة ، فما هو احتمال أن نحوي هذه المجموعة الصغيرة أكثر من 6 عناصر معيبة ؟

#### الحل

لنرمز بـ  $Y$  لعدد العناصر المعيبة الموجودة في العينة المستخرجة من البضاعة . نلاحظ أن للمتغير  $Y$  توزيعاً حدانياً بالمتغيرين  $n = 100$  ،  $P = 0.05$  . بما أن حجم العينة  $n$  كبيراً ، فإننا نستخدم التقريب الطبيعي الحداني حيث نأخذ :

$$\mu = n.p = 100 . (0.05) = 5$$

$$\sigma = \sqrt{np.q} = \sqrt{(100)(0.05)(0.95)} = 2.179$$

ومنه :

$$P[Y > 6] = P\left[\frac{Y - \mu}{\sigma} > \frac{6 - 5}{2.179}\right] = P[Z > 0.458]$$

$$= 1 - P[Z < 0.458]$$

$$= 1 - 0.6772$$

$$= 0.3228$$

وقد يسأل سائل ، أما كان بالإمكان حل هذه المسألة بطريقة الجدول II . والجواب يبدو واضحاً إذا لاحظنا :

$$P[Y > 6] = \sum_{y=7}^{100} b(y; 100, 0.5)$$

إذ إن الجدول II صمم من أجل قيم  $n$  من الصفر وحتى  $n = 20$  . ومن المتعذر استخدام الجدول المذكور نظراً لضخامة العدد  $n = 100$  .

#### مثال (٤,١٠)

في أحد اختبارات الرياضيات طرح على سعيد 250 سؤالاً مكتوباً ، وأمام كل سؤال أربع إجابات واحدة منها تمثل الإجابة الصحيحة ، واختار سعيد 100 سؤالاً منها ، وأجاب عليها بالتحزير ( التخمين ) نظراً لعدم تحضيره لهذه المادة بالشكل الجيد . ما هو احتمال أن يقود هذا التخمين إلى إجابات صحيحة عددها من 35 إلى 40 ؟

## الحل

نفرض أن  $X$  يمثل عدد الإجابات الصحيحة بين 100 إجابة ، فيكون المطلوب حساب :

$$P[35 \leq X \leq 40] = \sum_{x=35}^{40} b\left(x; 100, \frac{1}{4}\right)$$

وباستخدام التقريب الطبيعي للتوزيع الحداني بالوسيطين :

$$\mu = n.p = (100) \frac{1}{4} = 25$$

$$\sigma = \sqrt{npq} = \sqrt{(100) \left(\frac{1}{4}\right) \left(\frac{3}{4}\right)} = 4.629$$

نجد أن المساحة الواقعة بين النقطتين  $x_1 = 34.5$  ،  $x_2 = 40.5$  توافق المساحة بين القيمتين التاليتين للمتغير  $Z$  :

$$Z_1 = \frac{34.5 - 25}{4.629} = 2.05$$

$$Z_2 = \frac{40.5 - 25}{4.629} = 3.348$$

وهكذا نجد أن :

$$\begin{aligned} P[35 \leq X \leq 40] &\approx P[2.05 < Z < 3.348] \\ &= P[Z < 3.35] - P[Z < 2.05] \end{aligned}$$

ومن الجدول IV نجد أن :

$$P[25 \leq X \leq 30] = 0.9996 - 0.9798 \\ = 0.0198$$

### (٤,٤) التوزيعات غما ، الأسّي ، وكاي - مربع

#### Gamma, exponential and chi-Square distribution

على الرغم من أن التوزيع الطبيعي يمكن أن يستخدم لحل الكثير من المسائل الهندسية والعلمية ، فإن هناك عدداً من الحالات تتطلب نوعاً آخر من دوال الكثافة . من هذ الدوال ، هناك ما يسمى بالتوزيع غماً ، التوزيع الأسّي ، وأيضا التوزيع كاي مربع . سنبدأ فيما يلي بدراسة التوزيع غما .

لقد اشتق التوزيع غما اسمه من دالة معروفة بشكّل جيد لدى الرياضيين تدعى بالدالة غما ، وقبل دراسة هذا التوزيع ، لابد من ذكر بعض الخواص الهامة لهذه الدالة .

#### تعريف (٤,٢) الدالة غما Gamma function

تعرف الدالة غما بأنها :

$$\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} \cdot e^{-x} \cdot dx \quad : \quad x > 0$$

والخواص التي نحققها هذه الدالة نوردّها على التّالي :

- ١ —  $\Gamma(\alpha) = (\alpha - 1) \Gamma(\alpha - 1)$
- ٢ —  $\Gamma(\alpha) = (\alpha - 1) !$  عندما تكون  $\alpha$  رقماً صحيحاً موجباً
- ٣ —  $\Gamma(1) = 1$
- ٤ —  $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$

ويمكن بسهولة برهان هذه الخواص . فمثلاً لو كُاملنا بالتجزئة التكامل الوارد في التعريف (٤,٢) بأن نضع :  $u = x^{\alpha-1}$  ,  $dv = e^{-x} dx$  فإننا نجد أن :

$$\begin{aligned}\Gamma(\alpha) &= -e^{-x} \cdot \frac{x^{-1}}{0} \Big|_0^{\infty} + \int_0^{\infty} e^{-x} (\alpha-1) \cdot x^{-2} dx \\ &= (\alpha-1) \int_0^{\infty} e^{-x} \cdot x^{-2} dx\end{aligned}$$

$$\Gamma(\alpha) = (\alpha-1) \Gamma(\alpha-1)$$

وهو برهان الخاص الأول . ولو أعدنا مع التكرار ما ذكرناه لوجدنا أن :

$$\begin{aligned}\Gamma(\alpha) &= (\alpha-1) \cdot (\alpha-2) \Gamma(\alpha-2) \\ &= (\alpha-1) (\alpha-2) \cdot (\alpha-3) \Gamma(\alpha-3) \\ &= (\alpha-1) (\alpha-2) \dots \Gamma(1)\end{aligned}$$

إلا أننا نلاحظ أن :

$$\Gamma(1) = \int_0^{\infty} e^{-x} dx = 1$$

وبذلك نحصل على :

$$\Gamma(\alpha) = (\alpha-1) !$$

وهكذا نتحقق الخاصتان ٢ ، ٣ أيضا . وبتمريض  $x = \frac{1}{2} u^2$  في التكامل :

$$\Gamma\left(\frac{1}{2}\right) = \int_0^{\infty} e^{-x} \cdot x^{-\frac{1}{2}} dx \quad \text{فإننا نجد أن :}$$

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{2} \int_0^{\infty} e^{-\frac{1}{2}u^2} du \quad \text{وكما نعلم فإن :}$$

$$\int_0^{\infty} e^{-\frac{1}{2}u^2} du = \frac{1}{2} \sqrt{2\pi}$$

وبذلك نعلم فإن :

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{2} \cdot \frac{1}{2} \sqrt{2\pi} = \sqrt{\pi}$$

وهو كما نرى برهان الخاص الرابع :

يتضمن تعريفنا للتوزيع غما الدالة غما .

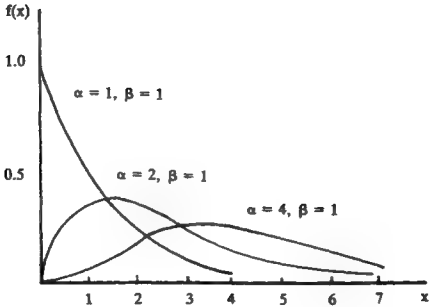
## التوزيع غاما Gamma distribution

نقول بأن للمتغير المستمر  $X$  توزيع غاما بالوسيطين  $\alpha, \beta$  إذا كانت كثافته من الشكل :

$$f(x) = \begin{cases} \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}} & : x > 0 \\ 0 & : \text{فيما عدا ذلك} \end{cases}$$

وبفرض أن كلا من  $\alpha, \beta > 0$ .

يوضح الشكل (٤، ١٤) عدداً من توزيعات غاما من أجل قيم متعددة لـ  $\alpha$  و  $\beta$ . وفي الحالة الخاصة إذا كان  $\alpha = 1$ ، فإننا نسمى التوزيع الناتج عن التوزيع غاما بالتوزيع الأسّي.



الشكل (٤، ١٤)

التوزيع غاما

## التوزيع الأسّي Exponential distribution

نقول بأن للمتغير المستمر  $X$  توزيعاً أسياً بالوسيط  $\beta$  إذا كانت دالة كثافته من الشكل

$$f(x) = \begin{cases} \frac{1}{\beta} \cdot e^{-\frac{x}{\beta}} & : x > 0 \\ 0 & : \text{فيما عدا ذلك} \end{cases}$$

حيث إن  $\beta > 0$  . وللتوزيع الأسّي تطبيقات واسعة في مجال الإحصاء وعلى الأخص في نظرية الثقة ، وفي مواضع الارتال . نسوق فيما يلي أحد هذه التطبيقات الممتعة لهذا التوزيع .

## مثال (٤,١١)

تحمى مجموعة نوعا خاصا من العناصر ، لكل منها عمر  $T$  ( يقدر السنوات ) ، ويتوزع وفقا للتوزيع الأسّي بالمتغير  $\beta = 6.5$  . أخذت عشرة عناصر منها وركبت ( نصبت ) بأشكال مختلفة ، ما هو احتمال أن يستمر سبعة منها على الأقل في تأدية وظائفها بعد مرور تسعة أعوام على تركيبها ؟

## الحل

نلاحظ أن احتمال أن يستمر عنصر ما من هذه العناصر في تأدية وظيفته بعد مرور تسعة أعوام على تركيبه هو :

$$P(T > 9) = \frac{1}{6.5} \int_9^{\infty} e^{-\frac{t}{6.5}} dt = e^{-\frac{9}{6.5}} = 0.25$$



لنفرض أن  $X$  يمثل عدد العناصر التي استمرت في تأدية وظائفها بعد مرور تسعة أعوام على تركيبها ، عندئذ باستخدام التوزيع الحداني نجد أن :

$$\begin{aligned} P(X \geq 7) &= \sum_{x=7}^{10} b(x; 10, 0.25) \\ &= 1 - \sum_{x=0}^6 b(x; 10, 0.25) \\ &= 1 - 0.9965 \\ &= 0.0035 \end{aligned}$$

يدعى التوزيع الخاص الثاني الذي نحصل عليه من توزيع غما بوضع  $\alpha = \frac{\nu}{2}$  ,  $\beta = 2$  ( حيث يمثل  $\nu$  عدداً صحيحاً موجباً ) بالتوزيع كاي مربع بـ  $\nu$  درجة من الحرية .

### التوزيع كاي - مربع Chi-Square distribution

يتوزع المتغير العشوائي المستمر  $X$  وفقاً للتوزيع غما بـ  $\nu$  درجة من الحرية إذا كانت دالة كثافته من الشكل :

$$f(x) = \begin{cases} \frac{1}{2^{\frac{\nu}{2}} \Gamma(\frac{\nu}{2})} x^{\frac{\nu}{2}-1} \cdot e^{-\frac{x}{2}} : x > 0 \\ 0 : \text{فيما عدا ذلك} \end{cases}$$

حيث يمثل  $\nu$  عدداً صحيحاً موجباً .

ويعتبر توزيع غما واحداً من الأدوات الرئيسية في حقل اختبار الفرضيات الذي سندرسه في الفصول القادمة . سنحاول فيما يلي إيجاد توقع وتباين كل من هذه المتغيرات العشوائية .

### نظرية (٤,٣)

إن توقع وتباين متغير واييل يعطيان بالعلاقتين :

$$\mu = \alpha \cdot \beta, \sigma^2 = \alpha \cdot \beta^2$$

### البرهان

نلاحظ أن العزم من المرتبة  $r$  حول المبدأ نكتب بالشكل التالي :

$$\mu_r = E(X^r) = \frac{1}{\beta^\alpha \cdot \Gamma(\alpha)} \int_0^\infty x^{r+\alpha-1} \cdot e^{-\frac{x}{\beta}} dx$$

وبإجراء تغير في المتحول بواسطة العلاقة  $\frac{x}{\beta} = y$ . نجد أن :

$$\begin{aligned} \mu_r &= \frac{\beta^r}{\Gamma(\alpha)} \int_0^\infty y^{\alpha+r-1} \cdot e^{-y} dy \\ &= \frac{\beta^r}{\Gamma(\alpha)} \Gamma(\alpha+r) \end{aligned}$$

وهكذا فإن :

$$\mu = \mu_1 = \frac{\beta \Gamma(\alpha+1)}{\Gamma(\alpha)} = \alpha \cdot \beta$$

كما نجد أيضا :

$$\begin{aligned} \sigma^2 &= \mu_2 - \mu^2 = \frac{\beta^2 \Gamma(\alpha+2)}{\Gamma(\alpha)} - \alpha^2 \cdot \beta^2 \\ &= \alpha \cdot \beta^2 \end{aligned}$$

نتيجة (٤,١)

إن توقع وتباين لتغير الأسى هما :

$$\mu = \beta, \sigma^2 = \beta^2$$

نتيجة (٤,٢)

أن توقع وتباين المتغير كاي - مربع هما :

$$\mu = \nu, \sigma^2 = 2 \cdot \nu$$

## (٤,٥) توزيع وايل Weibull distribution

تمكنت التكنولوجيا الحديثة من تصميم أنظمة عديدة معقدة تعتمد عملياتها على الوثوقية من العناصر المختلفة التي تتكون منها هذه الأنظمة . فمثلا يمكن للفيوز المستخدم في جهاز معين أن يحترق ، كما يمكن للعمود المعدني أن يلتوى ، ويمكن للجهاز حساس بالحرارة أن يتعطل . إن هذه العناصر المتماثلة والمعرضة لظروف بيئية معينة ستلاشى في أزمنة لا يمكن التنبؤ بها . إن عمر أى عنصر يقاس من فترة تكوينه إلى فترة فناءه ، وهذا الزمن هو متغير عشوائى مستمر  $T$  له كثافة احتمالية  $f(t)$  وإن أحد التوزيعات المستخدمة على نطاق واسع والذى يتعامل مع مثل هذه المشاكل كالوثوقية واختبار العمر يدعى بتوزيع وايل .

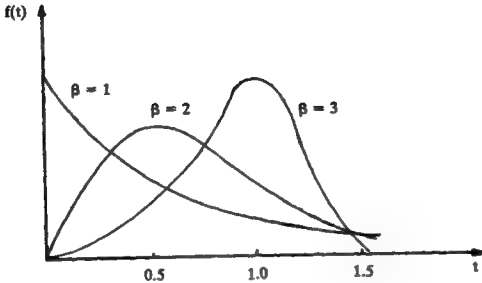
## توزيع وايل

المتغير العشوائى المستمر  $T$  له توزيع وايل بالمتغيرين  $\alpha$  ،  $\beta$  إذا كانت كثافته الاحتمالية من الشكل :

$$f(t) = \begin{cases} \alpha \cdot \beta \cdot t^{\beta-1} \cdot e^{-\alpha \cdot t^{\beta}} & : t > 0 \\ 0 & \text{فيما عدا ذلك:} \end{cases}$$

حيث  $\alpha > 0$  ,  $\beta > 0$

ولكثافة وايل منحنيات متعددة بحسب القيم المختلفة لكل من  $\alpha$  ,  $\beta$  . وفي الشكل (٤,٥) نرى أن منحنى الكثافة يتغير بشكل كبير بتغير قيم الوسطاء  $\alpha$  ,  $\beta$  وعلى الأخص بتغير قيم  $\beta$  . نلاحظ أن توزيع وايل يصبح توزيعاً أسياً إذا كان  $\beta = 1$  ومن أجل قيم  $\beta > 1$  يصبح شكل منحنى الكثافة قريباً من المنحنى الجرسى ويشبه المنحنى الطبيعي ، ويبدو عدم تناظره للعيان .



الشكل (٤,١٥)

## نظرية (٤,٣)

إن توقع وتباين متغير واييل يعطيان بالمعادلتين :

$$\mu = \alpha^{-\frac{1}{\beta}} \Gamma \left( 1 + \frac{1}{\beta} \right)$$

$$\sigma^2 = \alpha^{-\frac{2}{\beta}} \left\{ \Gamma \left( 1 + \frac{2}{\beta} \right) - \left[ \Gamma \left( 1 + \frac{1}{\beta} \right) \right]^2 \right\}$$

ولتطبيق توزيع واييل في نظرية الوثوقية ، نعرف أولاً وثوقية عنصر ، كاحتمال أن يكون دالة لزمن محدد على الأقل تحت ظروف تجريبية محددة . بفرض أن  $R(t)$  يمثل وثوقية عنصر ما في زمن  $t$  ، فإننا نستطيع أن نكتب عندئذ :

$$R(t) = P(X > t)$$

$$= \int_0^{\infty} f(t) dt$$

$$= 1 - F(t)$$

حيث يمثل  $F(t)$  التوزيع التراكمي للمتغير  $T$  . وإن الاحتمال الشرطي لفناء عنصر في مجال

زمنى من  $T = t$  إلى  $T = t + \Delta t$  يعطى بالعلاقة :

$$\frac{F(t + \Delta t) - F(t)}{R(t)}$$

وبتقسيم هذه النسبة على  $\Delta t$  ، ويأخذ النهاية عند ما تسعى  $\Delta t$  إلى الصفر ، فإننا نحصل على معدل الفشل ، والذي نرمز له بالرمز  $Z(t)$  . لذلك فإن :

$$\begin{aligned} Z(t) &= \lim_{\Delta t \rightarrow 0} \frac{F(t + \Delta t) - F(t)}{\Delta t} \cdot \frac{1}{R(t)} \\ &= \frac{F'(t)}{R(t)} = \frac{f(t)}{R(t)} = \frac{f(t)}{1 - F(t)} \end{aligned}$$

والعلاقة :

$$Z(t) = \frac{f(t)}{1 - F(t)}$$

تعبّر عن معدل الفشل في صيغة توزيع زمن الفشل .

وبما أن  $\bar{R}(t) = -\bar{F}(t)$  ،  $R(t) = 1 - F(t)$  فإننا نستطيع أن نكتب المعادلة التفاضلية التالية :

$$Z(t) = \frac{-\bar{R}(t)}{R(t)} = \frac{-d [\ln R(t)]}{dt}$$

وبحلها نجد أن :

$$\ln R(t) = - \int Z(t) dt$$

أو :

$$R(t) = e^{- \int Z(t) dt} + C$$

مثال (١٢، ٤)

بين أن معدل الفشل يعطى بالعلاقة التالية :

$$Z(t) = a \cdot b \cdot t^{b-1} : t > 0$$

إذا فقط إذا كان زمن الفشل يتوزع وفق توزيع واييل

$$f(t) = \alpha \cdot \beta \cdot t^{\beta-1} \cdot e^{-\alpha \cdot t^{\beta}} : t > 0$$

الحل

الحالة الأولى :

نفرض أن :

$$Z(t) = \alpha \cdot \beta \cdot t^{\beta-1} : t > 0$$

نلاحظ أن :

$$R(t) = \frac{-\int z(t)dt}{e^{-\alpha \cdot t^{\beta}}} = \frac{-\int \alpha \cdot \beta \cdot t^{\beta-1} dt}{e^{-\alpha \cdot t^{\beta}}} = \frac{-\alpha \cdot t^{\beta}}{e^{-\alpha \cdot t^{\beta}}} + C$$

ومن كون أن  $R(0) = 1$  ، نجد أن  $C = 0$  لذلك فإن :

$$R(t) = \frac{-\alpha \cdot t^{\beta}}{e^{-\alpha \cdot t^{\beta}}} : t > 0$$

كما نجد أيضا أن :

الحالة الثانية :

نفرض أن :

$$f(t) = \alpha \cdot \beta \cdot t^{\beta-1} \cdot e^{-\alpha \cdot t^{\beta}} : t > 0$$

نعلم أن :

$$Z(t) = \frac{f(t)}{R(t)}$$

حيث يمثل :

$$\begin{aligned} R(t) &= 1 - F(t) = 1 - \int_0^t \alpha \cdot \beta \cdot x^{\beta-1} \cdot e^{-\alpha \cdot x^{\beta}} dx \\ &= 1 + \int_0^t d(e^{-\alpha \cdot x^{\beta}}) \\ &= e^{-\alpha \cdot t^{\beta}} \end{aligned}$$

$$Z(t) = \frac{\alpha \cdot \beta \cdot t^{\beta-1} \cdot e^{-\alpha \cdot t^{\beta}}}{e^{-\alpha \cdot t^{\beta}}} = \alpha \cdot \beta \cdot t^{\beta-1} : t > 0$$

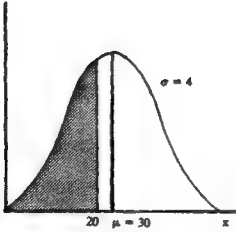
وهكذا نجد أن :

## تمارين محلولة

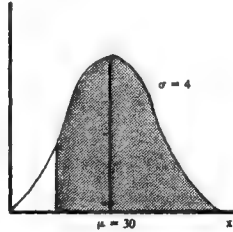
## تمرين (١)

بفرض  $X$  متغير عشوائى طبيعى بالمتغيرين  $\mu = 30, \sigma = 4$  . أوجد :

- (١) المساحة الواقعة أسفل النقطة 20 .
- (٢) المساحة الواقعة أعلى النقطة 25 .
- (٣) المساحة الواقعة بين النقطتين 33, 42 .
- (٤) النقطة التى تحضر قبلها مساحة قدرها 45% .
- (٥) النقطة التى تحدد بعدها مساحة قدرها 13% .



الشكل (١)



الشكل (٢)

## الحل

من الشكل (١) نلاحظ أن المساحة الواقعة أسفل النقطة 20 تمثل مساحة القسم المظلل من الشكل لذلك فإن :

$$\begin{aligned}
 P [X < 20] &= P \left[ \frac{X - \mu}{\sigma} < \frac{20 - 30}{4} \right] \\
 &= P \left[ Z < -\frac{10}{4} \right] \\
 &= P [Z < -2.5]
 \end{aligned}$$

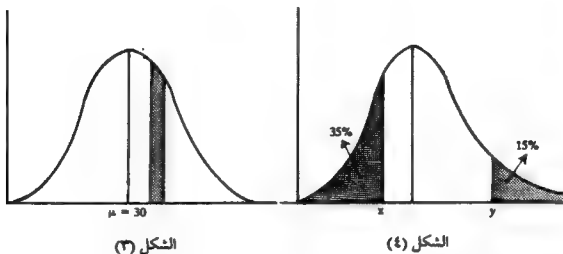
حيث يمثل  $Z$  متغيراً طبيعياً معيارياً . وبالعودة إلى الجدول IV نجد أن :

$$P [X < 20] = 0.0062$$

كذلك فإن المساحة الواقعة أعلى النقطة 25 تمثل مساحة القسم المظلل على الشكل (٧) وهذه المساحة تحسب بالعلاقة التالية :

$$\begin{aligned}
 P [X < 25] &= 1 - P[X > 25] \\
 &= 1 - P \left[ Z < -\frac{25-30}{4} \right] \\
 &= 1 - P \left[ Z < -\frac{-5}{4} \right] \\
 &= 1 - P [Z < -1.25] \\
 &= 1 - 0.1056 \\
 &= 0.8944
 \end{aligned}$$





يوضح القسم المظلل على الشكل (٣) المساحة المحصورة بين النقطتين 33 ، 42 ، وهذه المساحة تحسب على النحو التالي :

$$P [33 < X < 42] = P \left[ \frac{33-30}{4} < Z < \frac{42-30}{4} \right]$$

$$= P [0.75 < Z < 3]$$

$$= P[Z < 3] - P[Z < 0.75]$$

$$= 0.9987 - 0.7734$$

$$= 0.2253$$

كما أن النقطة x التي تحصر قبلها مساحة قدرها 35% موضحة على الشكل (٤) . يمثل القسم المظلل على الشكل المساحة المذكورة .

رابعاً : نجد أن النقطة التي تحصر قبلها مساحة قدرها 35% هي النقطة x المحققة للعلاقة :

$$P[X < x] = 0.35$$

ومنه :

$$P \left[ Z < \frac{x - 30}{4} \right] = 0.35$$

ونلاحظ من الجدول IV أنه يقابل المساحة 0.35 النقطة - 0.385  
ولذلك فإن :

$$x = (- 0.385) (4) + 30$$

ومنه :

$$x = 28.46$$

وأخيراً فإن النقطة التي تحدد بعدها مساحة قدرها 15% هي النقطة x المحققة للعلاقة :

$$P [X > x] = 0.15$$

$$P \left[ Z > \frac{x - 30}{4} \right] = 0.15$$

ومنه :

$$1 - P \left[ Z < \frac{x - 30}{4} \right] = 0.15$$

ولذلك فإن :

$$P \left[ Z < \frac{x - 30}{4} \right] = 1 - 0.15$$

ومن الجدول IV نجد أن النقطة التي تحدد لنا مساحة قدرها 0.85 هي النقطة 1.035  
ومنه :

$$Z = \frac{x - 30}{4} = 1.035$$

ولذلك فإن :

$$x = 4 \cdot (1.035) + 30$$

$$x = 34.14$$

تمرين (٢)

$Y$  متغير عشوائى طبيعى بالوسيطين  $\mu = 18, \sigma = 2.5$  . حدد القيمة  $k$  بحيث تتحقق المساواة :

$$P[X > k] = 0.1539$$

الحل

نلاحظ أن :

$$P \left[ \frac{X - \mu}{\sigma} > \frac{k - 18}{2.5} \right] = 0.1539$$

$$P \left[ Z > \frac{k - 18}{2.5} \right] = 0.1539$$

ومنه :

$$1 - P \left[ Z < \frac{k - 18}{2.5} \right] = 0.1539$$

ولذلك فإن :

$$P \left[ Z < \frac{k - 18}{2.5} \right] = 1 - 0.1539 = 0.8461$$

هذا ، ويقابل المساحة 0.8461 فى الجدول IV النقطة :

$$Z = \frac{k - 18}{2.5} = 1.02$$

ولذلك فإن :

$$k = 20.55$$

تمرين (٣)

أوجد الخطأ المرتكب في حساب  $\sum_{x=1}^4 b(x; 20, 0.1)$  بواسطة التقريب بالتوزيع الطبيعي .

الحل

من الواضح أن :

$$\sum_{x=1}^4 b(x; 20, 0.1) = \sum_{x=0}^4 b(x; 20, 0.1) - b(0; 20, 0.1)$$

من الجدول II نجد أن :

$$= 0.9568 - 0.1216$$

$$= 0.8352$$

وحسب التقريب بالتوزيع الطبيعي نجد أن :

$$\sum_{x=1}^4 b(x; 20, 0.1) = P[1 \leq X \leq 4]$$

ونلاحظ أن :

$$n = 20, p = 0.1$$

ولذلك فإن :

$$\sigma = \sqrt{np \cdot q} = 1.34, \quad \mu = np = 2$$

كما أن :

$$P[1 \leq X \leq 4] \approx P[0.5 < X < 4.5]$$

$$\begin{aligned}
&= P \left[ \frac{0.5 - 2}{1.34} < Z < \frac{4.5 - 2}{34} \right] \\
&= P \left[ Z < \frac{4.5}{1.34} \right] - P \left[ Z < \frac{-1.5}{1.34} \right] \\
&= P[Z < 1.869] - P[Z < -1.119] \\
&= 0.9686 - 0.1314 = 0.8372
\end{aligned}$$

والخطأ المرتكب هو :

$$0.8373 - 0.8352 = 0.0021$$

تمرين (٤)

إذا علمت أن احتمال شفاء إنسان من مرض القلب نتيجة إجراء عملية جراحية له هو 0.9 . ما هو احتمال شفاء من 84 إلى 95 مريضاً من أصل مئة أجريت لهم نفس العملية ؟

الحل

نفرض أن  $X$  يمثل عدد المرضى الذين تم شفاؤهم من أصل مئة مريض . نلاحظ أن للمتغير  $X$  توزيعاً حدانياً في تجربة فيها  $n = 100$  ،  $p = 0.9$  ،  $p = 90$  ،  $\mu = n \cdot p = 90$  ،  $\sigma = 3$  . نجد أن :

$$\begin{aligned}
P[84 \leq X \leq 95] &= \sum_{x=84}^{95} b(x; 100, 0.9) \\
&= P[83.5 < X < 95.5] \\
&= P \left[ \frac{83.5 - 90}{3} < Z < \frac{95.5 - 90}{3} \right]
\end{aligned}$$

$$= P[Z < 1.83] - P[Z < -2.166]$$

$$= 0.9664 - 0.0150$$

$$= 0.9514$$

تمرين (٥)

متغير عشوائي  $X$  له توزيع غما بالمتغيرين  $\alpha = 2$ ,  $\beta = 1$  أوجد :

$$P[1.8 < X < 2.4]$$

الحل

من المعلوم أن الكثافة الاحتمالية للمتغير  $X$  في هذه الحالة هي :

$$f(x) = \frac{1}{\Gamma^2 \cdot \Gamma(2)} x^{2-1} \cdot e^{-\frac{x}{1}} : x > 0$$

وبما أن  $\Gamma(2) = 1$  لذلك :

$$f(x) = \begin{cases} x \cdot e^{-x} \\ 0 \end{cases}$$

:  $x > 0$

فيما عدا ذلك :

وبما أن :

$$P[1.8 < X < 2.4] = \int_{1.8}^{2.4} f(x) dx$$

لذلك فإن :

$$= \int_{1.8}^{2.4} x \cdot e^{-x} dx$$

وبالتكامل بالتجزئة ، وبفرض أن  $U = x$  ،  $dv = e^{-x} dx$  نجد أن :

$$= - e^{-x} (x + 1) \Big|_{x=1.8}^{2.4}$$

$$= - 0.30844 + 0.462836$$

$$= 0.154396$$

تمرين (٦)

إذا علمت أن الاستهلاك اليومي للماء في مدينة معينة هو متغير عشوائي له تقريباً توزيع غما بالوسيطين  $\alpha = 2, \beta = 3$  ( علماً بأن هذا الاستهلاك بملايين اللترات ) . فإذا كان مخزون الماء اليومي في هذه المدينة 9 مليون لتر ، فما هو احتمال نفاذ الماء في أحد الأيام ؟

الحل

لنرمز للاستهلاك اليومي للماء في هذه المدينة بالرمز  $X$  ، فيكون لهذا المتغير توزيعاً قريباً من التوزيع غما بالوسيطين  $\alpha = 2, \beta = 3$  وهذا يعني أن :

$$f(x) = \begin{cases} \frac{1}{3^2 \cdot \Gamma(2)} x^{2-1} e^{-\frac{x}{3}} & : x > 0 \\ 0 & : \text{فيما عدا ذلك} \end{cases}$$

والاحتمال المطلوب هو :

$$P[X \geq 9] = 1 - P[X < 9]$$

$$= 1 - \int_0^9 f(x) dx$$

$$= 1 - \int_0^9 \frac{1}{9} x \cdot e^{-\frac{x}{3}} dx$$

وبإجراء تغيير في المتحول بالعلاقة  $U = \frac{x}{3}$  نجد أن :

$$P [ x \geq 9 ] = 1 - \int_0^3 U e^{-U} dU$$

وبالمكاملة بالتجزئة نجد أن :

$$\begin{aligned} &= 1 + \int_0^3 U d(e^{-U}) \\ &= 1 + Ue^{-U} \Big|_0^3 - \int_0^3 e^{-U} dU \\ &= 1 + 3e^{-3} + \left[ e^{-U} \right]_0^3 \\ &= 1 + 3e^{-3} + e^{-3} - 1 \\ &= 4e^{-3} \\ &= 0.1992 \end{aligned}$$

تموين (٧)

إذا كان عمر نوع معين من الأضرار الكهربائية ( بالأعوام ) يتوزع وفقاً للتوزيع الأسى . بمعدل فشل قدره  $\beta = 2$  . ركبنا مئة زر كهربائي من هذه الأضرار بأشكال مختلفة . ما هو احتمال أن يتعطل على الأكثر ثلاثين زراً من هذا الأضرار خلال العام الأول ؟

الحل

نرمز لعدد الأضرار المعطلة من بين المئة المركبة بـ  $X$  .



كما نرسم به  $Y$  لعمر كل زر بالسنوات

نلاحظ أن لـ  $Y$  كثافة احتمالية قدرها :

$$f(y) = \begin{cases} \frac{1}{2} e^{-\frac{y}{2}} & : y > 0 \\ 0 & \text{فيما عدا ذلك} \end{cases}$$

أما احتمال أن يتعطل أى زر خلال عام من تركيبه فهو يساوى :

$$P[Y \leq 1] = \int_0^1 \frac{1}{2} e^{-\frac{y}{2}} dy \quad \text{ومنه :}$$

$$p = 0.3935$$

$$q = 0.6065 \quad \text{وذلك فإن :}$$

وهكذا نجد أن المتوسط والانحراف المعياري للمتغير  $X$  هما :

$$\mu = n.p = 39.34, \sigma = 4.885$$

وهكذا نجد أن الاحتمال المطلوب هو :

$$P[X \leq 30] = P[X < 30.5]$$

$$= P \left[ Z < \frac{30.5 - 39.34}{4.885} \right]$$

$$= P[Z < -1.81]$$

من الجدول IV نجد أن :

$$= 0.0352$$

## تمارين عامة

(١) متغير عشوائى طبيعى وسطاؤه  $\mu = 200$  و  $\sigma^2 = 100$  أو جد :

أ - المساحة الواقعة تحت المنحنى الطبيعى وأعلى النقطة 179

ب - المساحة الواقعة تحت المنحنى الطبيعى وأسفل النقطة 214

(٢) باستخدام جدول التوزيع الطبيعى احسب الاحتمالات التالية :

$$P[-1 < Z < 1], P[-0.9 < Z < 0], P[Z > -0.9]$$

(٣) يتوزع عمر نوع من التلفزيونات وفقا للتوزيع الطبيعى بالوسط  $\mu = 3.1$

سنة والانحراف المعيارى 1.2 سنة . فإذا كانت الشركة تكفل كل تلفزيون لمدة سنة .

فما هى نسبة التلفزيونات المباعة التى ستضطر الشركة لاستبدالها بتلفزيون جديد ؟

(٤) إذا علمت أن عدد الصفقات التجارية التى يقوم بها التاجر محمد يتبع

التوزيع الحدائى ، وإذا علمت أن احتمال اتمام صفقة مع زبون عند قدومه هو 0.4 . فما

هو احتمال ألا يقل عدد صفقاته التى يتمها عن العدد 25 صفقة إذا استقبل مئة زبون فى

يوم معين ؟

(٥) استخدم جدول  $\chi^2$  لحساب الاحتمالات التالية :

$$أ - P[\chi^2 \geq 30.578] \text{ من أجل } \nu = 15$$

$$ب - P[\chi^2 \geq 130.844] \text{ من أجل } \nu = 26$$

$$ج - P[\chi^2 < 0.484] \text{ من أجل } \nu = 4$$

$$د - P[\chi^2 < 20.278] \text{ من أجل } \nu = 7$$

$$هـ - P[4.575 < \chi^2 < 24.725] \text{ من أجل } \nu = 11$$

(٦) إذا علمت أن جملة من الملاحظات تتوزع وفقا للتوزيع الطبيعى . ما هى

النسبة المئوية لاختلاف هذه الملاحظات عن المتوسط بـ

$$أ - أكثر من  $1.3\sigma$$$

$$ب - أقل من  $0.52\sigma$$$

(٧) إذا علمت أن معدل سقوط الأمطار في مدينة ما في شهر مارس ( آذار ) هو 9.22 سم مقرباً لأقرب مئة من الستيمتر . وبفرض أن سقوط الأمطار يتوزع وفقاً للتوزيع الطبيعي بالانحراف المعياري 2.82 سم . ما هو احتمال أن يهطل في مارس القادم في نفس المدينة .

- أ - أقل من 1.84 سم من المطر .  
 ب - أكثر من 5 سم وليس أكثر من 7 سم .  
 ج - أكثر من 13.8 سم .

(٨) احسب الخطأ المرتكب لدى تقريب  $\sum_{x=2}^{14} b(x; 8, 0.1)$  بواسطة المنحني الطبيعي .

(٩) أوجد التوقع والتباين لتوزيع واييل .

(١٠) نقول بأن للمتغير العشوائي  $X$  توزيعاً من نوع بتا بالنوسطين  $\alpha$  و  $\beta$  إذا كانت دالة كثافته من الشكل :

$$f(x) = \begin{cases} \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha) \cdot \Gamma(\beta)} x^{\alpha-1} \cdot (1-x)^{\beta-1} : 0 < x < 1 \\ 0 & \text{في بقية ذلك} \end{cases}$$

حيث إن :

$$\alpha > 0, \beta > 0$$

فإذا كانت النسبة المئوية للماركة معينة من التلفزيونات التي تتطلب التصليح خلال العام الأول لعملها تمثل متغيراً عشوائياً له توزيع بتا بالنوسطين ، فأوجد احتمال أن يكون 80% من التلفزيونات المباعة هذا العام من نفس الماركة تتطلب الإصلاح خلال السنة الأولى لإنتاجها .



# الفصل الخامس

## دوال المتغيرات العشوائية

- تغير المتغيرات
- الدوال المولدة للمزوم
- العينة العشوائية
- نظرية العينات
- توزيع المعاينة للوسط
- توزيع المعاينة للمتغير العشوائي
- $\frac{(n-1)S^2}{\sigma^2}$
- التوزيع t
- التوزيع F
- تقارين محلولة
- تقارين عامة .



## (٥, ١) تغيير المتغيرات Change of variables

نحتاج في أغلب الأحيان إلى معرفة التوزيع الاحتمالي لدالة بمتغير واحد أو أكثر . ففترض أن  $X$  يمثل ( متغيراً عشوائياً منقطعاً بالتوزيع الاحتمالي  $f(x)$  ، وأن  $Y = u(X)$  يمثل تطبيقاً واحداً لواحد بين مجموعة قيم  $X$  ومجموعة قيم  $Y$  ) . نحن نرغب في إيجاد التوزيع الاحتمالي للمتغير الجديد  $Y$  . ومن المهم أن نلاحظ أن التطبيق واحد لواحد الجديد يضمن لنا مقابلة كل قيمة  $x$  لـ  $X$  بقيمة وحيدة  $y$  لـ  $Y$  والعكس بالعكس .

ويبدو لنا واضحاً من خلال مناقشتنا للتوزيعات الاحتمالية المنقطعة في الفصل الثاني أن المتغير العشوائي  $Y$  يفترض القيمة  $y$  عندما يفترض المتغير  $X$  القيمة  $\lambda(y)$  حيث يمثل  $\lambda(y)$  الحل العكسي للمعادلة  $(Y = u(x))$  . وهكذا نجد أن التوزيع الاحتمالي للمتغير  $Y$  يعطى بالعلاقة :

$$h(y) = P(Y = y) = P[X = \lambda(y)] = f[\lambda(y)]$$

### نظرية (٥, ١)

نفرض أن  $X$  يمثل متغيراً عشوائياً منقطعاً بالتوزيع الاحتمالي  $f(x)$  ، وأن  $Y = u(x)$  يمثل تطبيقاً واحداً لواحد بين مجموعة قيم  $X$  ومجموعة قيم  $Y$  بحيث تكون للمعادلة  $Y = u(x)$  حلاً عكسياً من الشكل  $X = \lambda(y)$  . عندئذ يكون التوزيع الاحتمالي للمتغير  $Y$  :

$$h(y) = f[\lambda(y)]$$

### مثال (٥, ١)

نفرض أن  $X$  متغير عشوائي هندسي بالتوزيع الاحتمالي :

$$f(x) = \frac{2}{5} \cdot \left(\frac{3}{5}\right)^{x-1} ; x = 1, 2, 3, \dots$$

ولنفترض الآن توزيع الاحتمال للمتغير  $Y = x^3$

### الحل

بما أن جميع قيم المتغير  $X$  موجبة ، فالتطبيق  $Y$  هو تطبيق واحد لواحد بين مجموعة قيم  $X$  ومجموعة قيم  $Y$  . لذلك حسب النظرية (١, ٥) والعلاقة  $y = x^3$  التي تعطي  $X = Y^{1/3}$  نجد أن توزيع الاحتمال للمتغير  $Y$  هو :

$$h(y) = P\left[\sqrt[3]{y}\right] = \begin{cases} \left(\frac{2}{5}\right)\left(\frac{3}{5}\right)^{\frac{1}{3}-1} & : y = 1, 8, 27, \dots \\ 0 & : \text{فيما عدا ذلك} \end{cases}$$

لنفرض الآن لدينا متغيرين عشوائيين منقطعين  $X_1, X_2$  وتوزيعهما الاحتمالي المشترك هو  $f(x_1, x_2)$  . نحن نرغب في إيجاد التوزيع الاحتمالي المشترك  $g(y_1, y_2)$  لمتغيرين عشوائيين جديدين  $Y_1 = U_1(X_1, X_2), Y_2 = U_2(X_1, X_2)$  يحدد كلا منهما تقابلا واحدا لواحد بين مجموعة النقاط  $(x_1, x_2)$  ومجموعة النقاط  $(y_1, y_2)$  . بحل المعادلتين :

$$Y_2 = U_2(x_1, x_2), y_1 = U_1(x_1, x_2)$$

$$نجد أن :  $x_2 = \lambda_2(y_1, y_2), x_1 = \lambda_1(y_1, y_2)$  .$$

لذلك فإن المتغيرين  $Y_1, Y_2$  سيفرضان على الترتيب القيمتين  $y_1, y_2$  عندما يفترض المتغيران  $X_1, X_2$  القيمتين  $\lambda_1(y_1, y_2)$  و  $\lambda_2(y_1, y_2)$  على الترتيب أيضا . ولذلك فإن التوزيع الاحتمالي المشترك لـ  $Y_1, Y_2$  هو :

$$h(y_1, y_2) = P[Y_1 = y_1, Y_2 = y_2]$$

$$= P[X_1 = \lambda_1(y_1, y_2), X_2 = \lambda_2(y_1, y_2)]$$

$$= f[\lambda_1(y_1, y_2), \lambda_2(y_1, y_2)]$$



## نظرية (٥, ٢)

يفرض أن  $X_1, X_2$  متغيران عشوائيان منقطعان توزيعهما الاحتمالي المشترك  $f(y_1, y_2)$  وأن كلا من :  $Y_1 = U_1(X_1, X_2)$  و  $Y_2 = U_2(X_1, X_2)$  يمثلان تقابلاً واحداً لواحد بين مجموعة النقاط  $(x_1, x_2)$  و  $(y_1, y_2)$  بحيث يكون للمعادلتين  $Y_1 = U_1(y_1, y_2)$  و  $Y_2 = U_2(y_1, y_2)$  حلين وحيدين بالنسبة لـ  $x_1, x_2$  مثلاً  $x_1 = \lambda_1(y_1, y_2)$ ,  $x_2 = \lambda_2(y_1, y_2)$ . عندئذ يكون التوزيع الاحتمالي المشترك لـ  $Y_1, Y_2$  من الشكل :

$$h(y_1, y_2) = f[\lambda_1(y_1, y_2), \lambda_2(y_1, y_2)]$$

والنظرية (٥, ٢) مفيدة إلى أبعد حد في إيجاد التوزيع الاحتمالي لبض المتغيرات مثلاً ،  $Y_1 = U_1(X_1, X_2)$  ، وذلك في الحالة التي يكون فيها المتغيران  $X_1, X_2$  منقطعان ، وبالتوزيع الاحتمالي المشترك  $f(x_1, x_2)$  . ففي هذا نبحت ، قبل كل شيء ، عن توزيع الاحتمال المشترك  $h(y_1, y_2)$  . فيكون توزيع  $Y_1$  هو التوزيع الهامشي من الدالة  $h(y_1, y_2)$  ونجده بالجمع فوق جميع قيم  $y_2$  ، فإذا رمزنا لتوزيع  $Y_1$  بالرمز  $h(y_1)$  لوجدنا أن :

$$h(y_1) = \sum_{y_2} h(y_1, y_2)$$

## مثال (٥, ٢)

يفرض أن  $X_1, X_2$  يمثلان متغيرين عشوائيين مستقلان لهما توزيعين بواسونيين بالوسيطين  $\mu_1, \mu_2$  على الترتيب . لنبحث عن توزيع الاحتمال المشترك للمتغير الجديد  $Y_1 = X_1 + X_2$

## الحل

من الاستقلال العشوائى للمتغيرين  $X_1, X_2$  نجد أن :

$$f(x_1, x_2) = f(x_1) \cdot f(x_2)$$

ومنه :

$$f(x_1, x_2) = \frac{e^{-\mu_1} \cdot \mu_1^{x_1}}{x_1!} \cdot \frac{e^{-\mu_2} \cdot \mu_2^{x_2}}{x_2!}$$

$$= \frac{e^{-(\mu_1 + \mu_2)} \cdot \mu_1^{x_1} \cdot \mu_2^{x_2}}{x_1! \cdot x_2!}$$

حيث أن :

$$x_1 = 0, 1, 2, \dots, \quad x_2 = 0, 1, 2, \dots$$

لنعرف الآن متغيراً عشوائياً جديداً مثلاً  $Y_2 = X_2$  ، فتكون الحلول للمعادلتين  $Y_2 = X_2$  و  $Y_1 = X_1 + X_2$  من الشكل  $X_1 = Y_1 - Y_2$  ،  $X_2 = y_2$  وباستخدام النظرية (٥، ٢) نجد أن التوزيع الاحتمالي المشترك لكل من  $Y_1, Y_2$  هو :

$$h(y_1, y_2) = \frac{e^{-(\mu_1 + \mu_2)} \cdot \mu_1^{y_1 - y_2} \cdot \mu_2^{y_2}}{(y_1 - y_2)! \cdot y_2!}$$

حيث :

$$y_1 = 0, 1, 2, \dots, \quad y_2 = 0, 1, 2, \dots$$

وهكذا نجد أن التوزيع الهامشي لـ  $Y_1$  هو :

$$h(y_1) = \sum_{y_2=0}^{y_1} h(y_1, y_2)$$

$$= e^{-(\mu_1 + \mu_2)} \cdot \sum_{y_2=0}^{y_1} \frac{\mu_1^{y_1 - y_2} \cdot \mu_2^{y_2}}{(y_1 - y_2)! \cdot y_2!}$$

$$= \frac{e^{-(\mu_1 + \mu_2)}}{y_1!} \sum_{y_2=0}^{y_1} \frac{y_1!}{y_2! (y_1 - y_2)!} \mu_1^{y_1 - y_2} \cdot \mu_2^{y_2}$$

$$= \frac{e^{-(\mu_1 + \mu_2)}}{y_1!} \sum_{y_2=0}^{y_1} \binom{y_1}{y_2} \cdot \mu_1^{y_1 - y_2} \cdot \mu_2^{y_2}$$

وبملاحظة أن المجموع الأخير يمثل المقدار  $(\mu_1 + \mu_2)^{y_1}$  فإننا نجد أن :

$$h(y_1) = \frac{e^{-(\mu_1 + \mu_2)} (\mu_1 + \mu_2)^{y_1}}{y_1!} ; y_1 = 0, 1, 2, \dots$$

من هذا نستنتج أن مجموع متغيرين عشوائيين بواسونيين مستقلين بالوسيطين  $\mu_1, \mu_2$  هو من جديد متغير بواسوني وسيطه يساوى مجموع الوسيطين أى  $\mu_1 + \mu_2$  .

ولإيجاد التوزيع الاحتمالي للمتغير العشوائي  $Y = U(X)$  عندما يكون المتغير  $X$  مستمراً ، والتقابل  $U$  واحد لواحد فإننا نسوق النظرية التالية :

### نظرية (٥,٣)

ليكن  $X$  متغيراً عشوائياً مستمراً بالتوزيع الاحتمالي  $f(x)$  ، وليكن  $Y = U(X)$  تقابلاً واحداً لواحد بين مجموعة قيم  $X$  ومجموعة قيم  $Y$  بحيث يكون للمعادلة  $y = u(x)$  حلاً عكسياً وحيداً  $x = \lambda(y)$  . عندئذ يكون توزيع الاحتمال للمتغير  $Y$  من الشكل :

$$g(y) = f[\lambda(y)] \cdot |J|$$

حيث يمثل  $\lambda(y)$   $J$  ويدعى جاكوبى التقابل .

### البرهان :

#### الحالة الأولى

لنفرض أن  $y = U(x)$  يمثل دالة متزايدة كما هو موضح على الشكل (٥,١) . نلاحظ أنه إذا وقع  $Y$  في المجال  $(a, b)$  فعندئذ سيقع المتغير  $X$  في المجال  $(\lambda(a), \lambda(b))$  لذلك فإن :

$$P[a < Y < b] = P[\lambda(a) < X < \lambda(b)]$$

$$= \int_{\lambda(a)}^{\lambda(b)} f(x) dx$$

ويتغير المتحول في التكامل من  $x$  إلى  $y$  بالعلاقة  $x = \lambda(y)$  ، وبملاحظة أن  $dx = \lambda'(y) dy$  فإننا نجد أن :

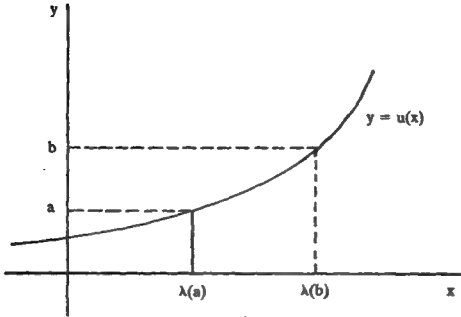
$$P[a < Y < b] = \int_a^b f[\lambda(y)] \cdot \lambda'(y) dy$$

ومن المعلوم أن :

$$P[a < Y < b] = \int_a^b h(y) dy$$

ومقارنة التكاملين السابقين نجد :

$$h(y) = f[\lambda(y)] \cdot \lambda'(y) = f[\lambda(y)] \cdot J$$



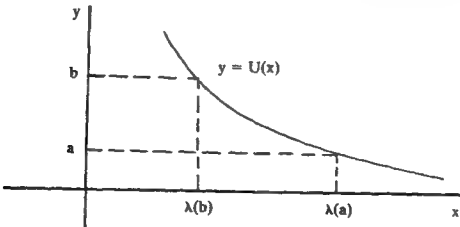
الشكل (٥,١)

وإذا لاحظنا أن  $J = \lambda'(y)$  يمثل ميل المماس لمنحنى دالة متزايدة  $y = U(X)$  ، لظهر لنا أن

$$h(y) = f[\lambda(y)] \cdot |J| \quad : \text{ولذلك فإن } J = |J|$$

الحالة الثانية

نفرض أن الدالة  $y = U(x)$  متناقصة كما هو موضح على الشكل (٥,٢)



الشكل (٥,٢)

عندئذ نلاحظ أن :

$$P[a < Y < b] = P[\lambda(b) < X < \lambda(a)]$$

$$= \int_{\lambda(b)}^{\lambda(a)} f(x) dx$$

وإذا غيرنا المتحول من  $x$  إلى  $y$  فإننا نجد أن :

$$\begin{aligned} P[a < Y < b] &= \int_{\lambda(b)}^{\lambda(a)} f[\lambda(y)] \cdot \lambda'(y) dy \\ &= - \int_a^b f[\lambda(y)] \cdot \lambda'(y) dy \end{aligned}$$

من هذا نستنتج أن :

$$h(y) = - f[\lambda(y)] \cdot \lambda'(y) = - f[\lambda(y)] \cdot J$$

وفي هذه الحالة نعلم أن ميل المماس سالب ومنه  $J = -|J|$  لذلك فإن :

$$h(y) = f[\lambda(y)] \cdot |J|$$

مثال (٥,٣)

بفرض  $X$  متغير عشوائي مستمر بالتوزيع الاحتمالي :

$$f(x) = \begin{cases} \frac{3}{7} x^2 & : 1 < x < 2 \\ 0 & : \text{فيما عدا ذلك} \end{cases}$$

ولنفترض عن توزيع الاحتمال للمتغير  $Y = X + 3$

الحل

إن الحل المكسبي للمعادلة  $y = x + 3$  هو  $y = x - 3$  ومنه نستنتج أن  $dx =$

dy . باستخدام النظرية (٥,٣) نجد أن الكثافة الاحتمالية للمتغير Y هي :

$$h(y) = \frac{3}{7} \cdot (y - 3)^2$$

ومنه :

$$h(y) = \begin{cases} \frac{3}{7} \cdot (y - 3)^2 & : +4 < y < 5 \\ 0 & : \text{فيما عدا ذلك} \end{cases}$$

لإيجاد التوزيع الاحتمالي المشترك للمتغيرين  $Y_2 = U_2(X_1, X_2)$ ,  $Y_1 = U_1(X_1, X_2)$  وذلك في الحالة التي يكون فيها المتغيران  $X_1, X_2$  مستمرين ، ويكون كل من التقابليين  $U_1, U_2$  ممثلاً لتقابل واحد لواحد ، فإننا نحتاج إلى النظرية الإضافية التالية والتي سنوردها بدون برهان .

#### نظرية (٥,٤)

بفرض أن  $X_1, X_2$  يمثلان متغيرين مستمرين بالتوزيع المشترك  $f(x_1, x_2)$  ، وأن كلا من  $Y_1 = U_1(X_1, X_2)$  ،  $Y_2 = U_2(X_1, X_2)$  يمثلان تقابلاً واحداً لواحد بين النقاط  $(x_1, x_2)$  و  $(y_1, y_2)$  بحيث يكون للمعادلتين  $Y_1 = U_1(x_1, x_2)$  و  $Y_2 = U_2(x_1, x_2)$  حلان وحيدان بالنسبة لـ  $x_1, x_2$  هما  $x_1 = \lambda_1(y_1, y_2)$  و  $x_2 = \lambda_2(y_1, y_2)$  عندئذ يكون التوزيع المشترك للمتغيرين العشوائيين  $Y_1, Y_2$  من الشكل :

$$h(y_1, y_2) = f[\lambda_1(y_1, y_2), \lambda_2(y_1, y_2)] \cdot |J|$$

حيث يمثل J جاكوبي الانتقال المحدد بالعلاقة :

$$J = \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} \end{vmatrix}$$

وحيث يمثل المشتق الجزئي لـ  $\lambda_i = w_i(y_1, y_2)$  بالنسبة لـ  $y_1$  على اعتبار أن  $y_2$  ثابتاً وأن  $i = 1, 2$ . وبنفس الطريقة نعرف بقية المشتقات الجزئية .

### مثال (٤, ٥)

بفرض أن  $x_1, x_2$  يمثلان متغيرين عشوائيين مستمران بالتوزيع الاحتمالي المشترك :

$$f(x_1, x_2) = \begin{cases} \frac{3}{32} x_1^2 x_2 & : 0 < x_1 < 2 \\ & 1 < x_2 < 3 \\ 0 & : \text{فيما عدا ذلك} \end{cases}$$

لنفتش عن التوزيع الاحتمالي المشترك للمتغيرين  $y_1 = x_1^3$  و  $y_2 = x_1 x_2$

### الحل

لنحل المعادلتين  $y_1 = x_1^3$  و  $y_2 = x_1 x_2$  بالنسبة لـ  $x_1, x_2$  فنجد أن :

$$x_1 = y_1^{1/3}, \quad x_2 = \frac{y_2}{y_1^{1/3}}$$

ومن هاتين المعادلتين نجد أن معين جاكوبي هو من الشكل :

$$\begin{vmatrix} \frac{1}{3y_1^{2/3}} & 0 \\ -\frac{y_2}{3y_1^{4/3}} & \frac{1}{y_1^{1/3}} \end{vmatrix} = \frac{1}{3y_1}$$

كما نلاحظ أن التقابلات تمثل واحداً لواحد بين مجموعة النقاط :

$$\{(x_1, x_2) | 0 < x_1 < 2, 1 < x_2 < 3\}$$

ومجموعة النقاط :

$$\{(y_1, y_2) \mid \frac{y_2^3}{27} < y_1 < 8, 0 < y_2 < 6\}$$

من النظرية (٥، ٤) نجد أن التوزيع الاحتمالي المشترك للمتغيرين  $Y_1, Y_2$  هو من الشكل :

$$h(y_1, y_2) = \frac{y_2}{32y_1}$$

أو

$$g(y_1, y_2) = \begin{cases} \frac{y_2}{32y_1} & : \frac{y_2^3}{27} < y_1 < 8 \\ 0 & : \text{فيما عدا ذلك} \end{cases}$$

تنحصر المشاكل الناشئة عن إيجاد توزيع الاحتمال للمتغير العشوائى  $Y = U(X)$  عندما يكون  $X$  مستمراً في الحالة التى لا يكون فيها التقابل واحداً لواحد بين مجموعة القيم  $x$  ومجموعة القيم  $y$ . ففي هذه الحالة نلاحظ أنه يقابل كل قيمة لـ  $y$  أكثر من قيمة واحدة لـ  $x$ . فمثلاً إذا كان  $f(x)$  موجبا فوق المجال  $-1 < x < 2$  ومعلوما فيما عدا ذلك ، وإذا اعتبرنا أن التقابل هو من الشكل  $Y = X^2$  ، فإننا نجد في هذه الحالة أن  $x = \pm \sqrt{y}$  من أجل  $0 < y < 1$  وأن  $x = \sqrt{y}$  من أجل  $1 < y < 4$ . وتوزيع الاحتمال لـ  $Y$  فوق المجال  $0 < y < 4$  نوجده باستخدام النظرية (٥، ٣)

$$g(y) = f[w(y)] |J| = \frac{f(\sqrt{y})}{2\sqrt{y}}, 1 < y < 4$$

وفي الحالة  $0 < y < 1$  يمكننا تجربة المجال  $-1 < x < 1$  لنحصل على الدالتين العكسيتين :

$$x = \begin{cases} -\sqrt{y} & : -1 < x < 0 \\ \sqrt{y} & : 0 < x < 1 \end{cases}$$



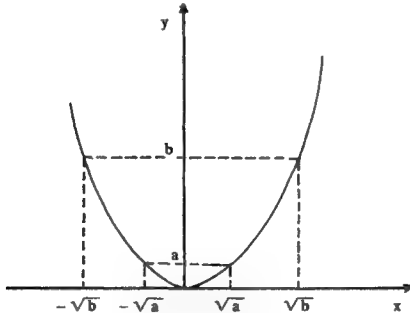
وهكذا نجد أن كل قيمة لـ  $y$  يقابلها قيمة وحيدة لـ  $x$  في كل مجال من مجالات التجزئة ، كما هو موضح على الشكل (٥,٣) . كما نلاحظ أن :

$$P[a < Y < b] = P[-\sqrt{b} < X < -\sqrt{a}] + P[\sqrt{a} < X < \sqrt{b}]$$

$$= \int_{-\sqrt{b}}^{-\sqrt{a}} f(x) dx + \int_{\sqrt{a}}^{\sqrt{b}} f(x) dx$$

وبتغير متحول التكامل من  $x$  إلى  $y$  نجد أن :

$$\begin{aligned} &= \int_b^a f(-\sqrt{y}) j_1 \cdot dy + \int_a^b f(\sqrt{y}) j_2 dy \\ &= - \int_a^b f(-\sqrt{y}) j_1 \cdot dy + \int_a^b f(\sqrt{y}) j_2 dy \end{aligned}$$



الشكل (٥,٣)

حيث أن :

$$J_1 = \frac{d(-\sqrt{y})}{dy} = \frac{-1}{2\sqrt{y}} = -|J_1|$$

$$J_2 = \frac{d(\sqrt{y})}{dy} = \frac{1}{2\sqrt{y}} = |J_2|$$

ولذلك يمكن أن نكتب :

$$P[a < Y < b] = \int_a^b (f(-\sqrt{y}) |J_1| + f(\sqrt{y}) \cdot |J_2|) dy$$

وهكذا نجد أن :

$$\begin{aligned} h(y) &= f(-\sqrt{y}) |J_1| + f(\sqrt{y}) \cdot |J_2| \\ &= \frac{[f(-\sqrt{y}) + f(\sqrt{y})]}{2\sqrt{y}}, \quad 0 < y < 1 \end{aligned}$$

وأخيراً فإن :

$$h(y) = \begin{cases} \frac{[f(-\sqrt{y}) + f(\sqrt{y})]}{2\sqrt{y}}, & 0 < y < 1 \\ \frac{f(\sqrt{y})}{2\sqrt{y}}, & 1 < y < 4 \\ 0, & \text{فيما عدا ذلك} \end{cases}$$

يمكن تعميم الإجراءات السابقة في البحث عن  $h(y)$  من أجل  $0 < y < 1$  بالنظرية (٥,٥) من أجل  $k$  دالة عكسية .

### نظرية (٥,٥)

لنفرض أن  $X$  يمثل متغيراً عشوائياً مستمراً له توزيع احتمال  $f(x)$  ، ولنفرض أن  $Y = U(X)$  يمثل تقابلاً ليس واحداً لواحد بين قيم  $X$  وقيم  $Y$  . فإذا أمكن تجزئة مجال تحول المتغير  $X$  إلى  $k$  مجالاً جزئياً بحيث تمثل كل دالة عكسية من الدوال :

$$x_1 = \lambda_1(y), \quad x_2 = \lambda_2(y), \quad \dots, \quad x_k = \lambda_k(y)$$

تقابلاً واحداً لواحد ، عندئذ يكون التوزيع الاحتمالي للمتغير  $Y$  من الشكل :

$$h(y) = \sum_{i=1}^k f[\lambda_i(y)] |J_i|$$

وحيث إن :

$$J_i = \lambda_i(y) \quad , \quad i = 1, 2, \dots, k$$

مثال (٥, ٥)

إذا كان للمتغير  $X$  توزيعاً طبيعياً بالوسيطين  $\mu$  ,  $\sigma$  , فعندئذ يكون للمتغير توزيع كاي - مربع بدرجة حرية واحدة .

الحل

نلاحظ أن للمتغير  $z = \frac{X - \mu}{\sigma}$  توزيعاً طبيعياً معيارياً وكثافة قدرها :

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$$

لنبحث الآن عن توزيع الاحتمال للمتغير  $Y = z^2$  . إن الحل المعكس للمعادلة  $y = z^2$  هو  $z = \pm \sqrt{y}$  . ل نرمز بـ  $z_1 = -\sqrt{y}$  ,  $z_2 = +\sqrt{y}$  . نلاحظ أن

$$J_1 = -\frac{1}{2\sqrt{y}} \text{ , } J_2 = \frac{1}{2\sqrt{y}}$$

لذلك حسب النظرية (٥, ٥) ولدينا :

$$\begin{aligned} g(y) &= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y} \left| \frac{-1}{2\sqrt{y}} \right| + \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y} \left| \frac{1}{2\sqrt{y}} \right| \\ &= \frac{1}{\sqrt{2\pi}} e^{-\frac{y}{2}} \cdot \frac{1}{y^{\frac{1}{2}} - 1} \quad , \quad y > 0 \end{aligned}$$

ومن المعلوم أن

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$$

لذلك فإن :

$$g(y) = \frac{1}{\sqrt{2\pi}} y^{\frac{1}{2} - 1} \cdot e^{-\frac{1}{2}y} \quad , \quad y > 0$$

$$2^{\frac{1}{2}} \Gamma\left(\frac{1}{2}\right)$$

وهذا يعنى أن للمتغير  $Y$  توزيعاً من نوع توزيع كاي - مربع بدرجة واحدة من الحرية .

### (٥,٢) الدوال المولدة للعزوم Moments generating functions

على الرغم من أن طريقة تغير المتغيرات تقودنا إلى طريقة فعالة في إيجاد توزيع دالة بعدة متغيرات ، فإن هناك أيضاً إجراءات مفضلة في الحالة التي تكون فيها المتغيرات مستقلة . وسوف نشير إلى هذه الإجراءات من خلال ما يسمى بالدوال المولدة للعزوم .

#### تعريف (٥,١) الدالة المولدة للعزوم

نعرف الدالة المولدة للعزوم لمتغير عشوائي  $X$  والتي نرمز لها بالرمز  $M_X(t)$  بإحدى العلاقتين التاليتين .

$$M_X(t) = E(e^{t \cdot X})$$

$$= \sum_x e^{tx} \cdot f(x) \quad : \quad \text{إذا كان } X \text{ منقطعاً}$$

$$= \int_{-\infty}^{+\infty} e^{tx} \cdot f(x) dx \quad : \quad \text{إذا كان } X \text{ مستمراً}$$

وتكون الدالة المولدة للعزوم موجودة إذا كان المجموع أو التكامل في التعريف (٥,١) محدوداً . فإذا كانت الدالة المولدة للعزوم موجودة ، فيمكن عندئذ استخدامها في إيجاد مختلف عزوم المتغير العشوائي ، وطريقة ذلك توضحها النظرية التالية .

#### نظرية (٥,٦)

إذا كان  $M_X(t)$  ممثلاً للدالة المولدة للعزوم للمتغير العشوائي  $X$  فعندئذ يكون :

$$\left. \frac{d^r M_X(t)}{dt^r} \right|_{t=0} = \mu_r'$$

## البرهان

من التعريف (٥,١) وباشتقاق ما تحت إشارة المجموع أو التكامل فإننا نحصل على ما يلي :

$$\frac{d^r M_x(x)}{dt^r} = \sum_{i=1}^n x_i^r \cdot e^{ix_i} f(x_i)$$

أو :

$$\frac{d^r M_x(x)}{dt^r} = \int_{-\infty}^{+\infty} x^r \cdot e^{ix} f(x) dx$$

وبوضع  $t = 0$  في كلا طرفي العلاقة السابقة فإننا نجد أن :

$$\mu_r' = E(X^r) = \left. \frac{d^r M_x(x)}{dt^r} \right|_{t=0}$$

## مثال (٥,٦)

لنبحث عن الدالة المولدة للعزوم للمتغير العشوائي الحداني  $X$  . من التعريف

$$M_x(t) = \sum_{x=0}^n e^{tx} \cdot \left( \frac{n}{x} \right) p^x \cdot q^{n-x} \quad (٥,١) \quad \text{نجد أن :}$$

$$= \sum_{x=0}^n \left( \frac{n}{x} \right) (p \cdot e^t)^x \cdot q^{n-x}$$

وهكذا نجد أن

$$M_x(t) = (p e^t + q)^n$$

وحسب النظرية (٥,٦) نجد بالاشتقاق مرتين أن :

$$\frac{dM_x(t)}{dt} = n \cdot p (p e^t + q)^{n-1} \cdot e^t$$

ومنه :

$$\mu = \mu_1' = \left. \frac{dM_x(t)}{dt} \right|_{t=0} = n \cdot p$$

وهو ما برهننا عليه سابقا . كذلك فإن :

$$\frac{d^2 M_x(t)}{dt^2} = n \cdot p [ e^t (p e^t + q)^{n-1} + p(n-1) (p e^t + q) e^{2t} ]$$

وبوضع  $t = 0$  نجد أن :

$$\mu_2 = n p [ 1 + p(n-1) ]$$

وهكذا نجد أن :

$$\sigma^2 = \mu_2 - \mu^2 = np + np^2 (n-1) - n^2 p^2$$

$$= np - np^2$$

$$= np (1 - p)$$

$$= np \cdot q$$

وهي نفس النتيجة التي وجدناها في الفصل الثالث .

### مثال (٥,٧)

لنبين أن الدالة المولدة للعزوم لمتغير عشوائى طبيعى بالتوزيع  $\mu$  والتباين  $\sigma^2$  هي من الشكل :

$$M_x(t) = e^{\mu t + \frac{1}{2} \sigma^2 \cdot t^2}$$

### الحل

بالحقيقة : من التعريف (٥,١) نجد أن :

$$\begin{aligned} M_x(t) &= \int_{-\infty}^{+\infty} e^{tx} \cdot \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot e^{-\frac{1}{2} \left( \frac{x-\mu}{\sigma} \right)^2} dx \\ &= \frac{1}{\sqrt{2\pi} \sigma} \int_{-\infty}^{+\infty} e^{-\frac{1}{2\sigma^2} [X^2 - 2\mu X + \mu^2 - 2\mu X]} dx \end{aligned}$$

وبإكمال المقدار  $x^2 - 2(\mu + t\sigma^2)x + \mu^2 = I$  إلى مربع كامل ، فإننا نجد أن :

$$I = [x - (\mu + t\sigma^2)]^2 - 2\mu t\sigma^2 - t^2 \sigma^4$$

وهكذا نجد أن :

$$M_x(t) = \frac{1}{\sqrt{2\pi}\sigma} e^{\mu t + \frac{1}{2} t^2 \sigma^2} \int_{-\infty}^{+\infty} e^{-\frac{1}{2} \left[ \frac{x - (\mu + t\sigma^2)}{\sigma} \right]^2} dx$$

وبإجراء تغيير في المتحول بالعلاقة  $\frac{x - (\mu + t\sigma^2)}{\sigma} = u$  نجد أن  $dx = \sigma du$

$$\begin{aligned} M_x(t) &= e^{\mu t + \frac{1}{2} t^2 \sigma^2} \cdot \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{1}{2} u^2} du \\ &= e^{\mu t + \frac{1}{2} t^2 \sigma^2} \end{aligned}$$

وهو المطلوب .

### مثال (٥,٨)

يفرض أن للمتغير  $X$  توزيعاً من نوع كاي - مربع بـ  $\nu$  درجة من الحرية ، فعندئذ تكون الدالة المولدة للعزوم لهذا المتغير من الشكل :

$$M_x(t) = \frac{1}{(1 - 2t)^{\nu/2}}$$

الحل

لقد حصلنا على توزيع كاي - مربع على شكل حالة خاصة من توزيع غما بوضع  $\beta = 2, \alpha = \frac{\nu}{2}$  . باستبدال قيمة  $f(x)$  في التعريف (٥,١) نجد أن :

$$\begin{aligned} M_x(t) &= \int_0^{\infty} e^{tx} \frac{1}{2^{\nu/2} \Gamma(\frac{\nu}{2})} x^{\frac{\nu}{2}-1} \cdot e^{-\frac{x}{2}} dx \\ &= \frac{1}{2^{\nu/2} \Gamma(\frac{\nu}{2})} \int_0^{\infty} x^{\frac{\nu}{2}-1} \cdot e^{-x(\frac{1}{2})} dx \end{aligned}$$

وبوضع  $\bar{y} = \left(\frac{1-2t}{2}\right) x$  و  $dx = \frac{2}{1-2t} dy$  ومنه :

$$M_x(t) = \frac{1}{2^{n/2} \left(\frac{\nu}{2}\right)} \int_0^\infty \left(\frac{2y}{1-2t}\right)^{\frac{\nu}{2}-1} \cdot e^{-y} \cdot \left(\frac{2}{1-2t}\right) dy$$

$$= \frac{1}{\Gamma\left(\frac{\nu}{2}\right) (1-2t)^{n/2}} \int_0^\infty y^{\frac{\nu}{2}-1} e^{-y} dy$$

وإذا لاحظنا أن :

$$\int_0^\infty y^{\frac{\nu}{2}-1} e^{-y} dy = \Gamma\left(\frac{\nu}{2}\right)$$

لوجدنا أن :

$$M_x(t) = \frac{1}{(1-2t)^{n/2}}$$

لنستعرض الآن أربع نظريات تمثل خواص الدوال المولدة للعزوم . هذه الخواص ستكون مفيدة جداً في إيجاد توزيع أى تركيب خطي لمجموعة متغيرات عشوائية مستقلة . أما النظرية الأولى (٥,٧) فسنوردها بلون برهان .

### نظرية (٥,٧) ( نظرية الوحدةانية )

إذا كان للمتغيرين العشوائيين  $X$  و  $Y$  نفس الدالة المولدة للعزوم ، فعندئذ يكون لهما نفس توزيع الاحتمال .

### نظرية (٥,٨)

$$M_{X+a}(t) = e^{at} \cdot M_X(t)$$

البرهان

$$M_{X+a}(t) = E[e^{t(X+a)}] = e^{at} E[e^{tX}] = e^{at} M_X(t)$$



## نظرية (٩، ٥)

$$M_{aX}(t) = M(a, t)$$

البرهان

$$\begin{aligned} M_{aX}(t) &= E \left[ e^{t(a)X} \right] = E \left[ e^{(ta)X} \right] \\ &= M_X(a \cdot t) \end{aligned}$$

## نظرية (١٠، ٥)

إذا كان المتغيران العشوائيان  $X_1, X_2$  مستقلين ، وكان لهما دالتان مولدتان للعزوم  $M_{X_1}(t)$  ،  $M_{X_2}(t)$  على الترتيب ، فعندئذ يكون للمتغير  $Y = X_1 + X_2$  دالة مولدة للعزوم من الشكل :

$$M_Y(t) = M_{X_1+X_2}(t) = M_{X_1}(t) \cdot M_{X_2}(t)$$

البرهان

$$\begin{aligned} M_Y(t) &= E(e^{ty}) = E \left[ e^{t(x_1 + x_2)} \right] \\ &= \int \int_{-\infty}^{+\infty} e^{t(x_1 + x_2)} f(x_1, x_2) dx_1 dx_2 \end{aligned}$$

وبما أن المتغيرين  $X_1, X_2$  مستقلان فهذا يعنى أن  $f(x_1, x_2) = g(x_1) h(x_2)$  ولذلك فإن :

$$\begin{aligned} M_Y(t) &= \int_{-\infty}^{+\infty} g(x_1) \frac{e^{tx_1}}{e} dx_1 \int_{-\infty}^{+\infty} h(x_2) \cdot \frac{e^{tx_2}}{e} dx_2 \\ &= M_{X_1}(t) \cdot M_{X_2}(t) \end{aligned}$$

ويم برهان النظرية السابقة من أجل عدة متغيرات منقطعة بنفس الطريقة وذلك بأخذ المجموع بدلا عن التكامل .

باستخدام النظريتين (٥,٧) ، (٥,١٠) مع الانتباه إلى أن الدالة المولدة للعزوم

$$M_x(t) = e^{\mu(e^t - 1)}$$

لتغير عشوائى بواسونى هى من الشكل

نجد أنه إذا كان  $X_1, X_2$  متغيرين من النوع البواسونى وكانا مستقلين ، وإذا كان الدالتان المولدتان للعزوم لهما على الترتيب  $M_{X_1}(t) = e^{\mu_1(e^t - 1)}$  ،  $M_{X_2}(t) = e^{\mu_2(e^t - 1)}$  فعندئذ تكون الدالة المولدة للعزوم للمتغير  $Y = X_1 + X_2$  من الشكل :

$$\begin{aligned} M_Y(t) &= M_{X_1}(t) \cdot M_{X_2}(t) \\ &= e^{\mu_1(e^t - 1)} \cdot e^{\mu_2(e^t - 1)} \\ &= e^{(\mu_1 + \mu_2)(e^t - 1)} \end{aligned}$$

وهذا يعنى أن مجموع متغيرين بواسونيين ( بالتوقعين  $\mu_1, \mu_2$  ) هو متغير بواسونى وسيطه يساوى مجموع الوسيطين  $\mu_1 + \mu_2$  .

وغالبا ما نحتاج فى التطبيقات الإحصائية إلى تعريف التوزيع الاحتمالى لأى تركيب خطى بمتغيرات عشوائية طبيعية . لنفتش عن التوزيع الاحتمالى للمتغير  $Y = a_1 X_1 + a_2 X_2$  إذا كان  $X_i$  طبيعيا بالتوقع  $\mu_i$  والتباين  $\sigma_i^2$  حيث  $i = 1, 2$  ، وهذه المتغيرات مستقلة أيضا . نلاحظ أولا حسب النظرية (٥,١٠) أن

$$M_Y(t) = M_{a_1 X_1}(t) \cdot M_{a_2 X_2}(t)$$

وباستخدام النظرية (٥,٩) نجد أن :

$$= M_{X_1}(a_1 t) \cdot M_{X_2}(a_2 t)$$

وباستخدام عبارة الدالة المولدة للعزوم لمتغير طبيعى والتى استخرجناها فى المثال (٥,٧) نجد أن :

$$= \frac{a_1^2 \mu_1 + a_1^2 \sigma_1^2 / 2}{e} + \frac{a_2^2 \mu_2 + a_2^2 \sigma_2^2 / 2}{e}$$

$$= \frac{1}{e} (a_1 \mu_1 + a_2 \mu_2) + \frac{1}{2} e^2 (a_1^2 \sigma_1^2 + a_2^2 \sigma_2^2)$$

وبمقارنة العبارة الأخيرة بنتيجة المثال (٥،٧) نجد أن لـ  $Y$  توزيعاً طبيعياً بالوسيطين  $(a_1 \mu_1 + a_2 \mu_2, a_1^2 \sigma_1^2 + a_2^2 \sigma_2^2)$  ولتعميم هذه الحالة على  $n$  متغيراً طبيعياً نسوق النظرية التالية بدون برهان .

### نظرية (٥،١١)

إذا كانت المتغيرات  $X_1, X_2, \dots, X_n$  مستقلة وطبيعية بالتوقعات  $\mu_1, \dots, \mu_n$  والتباينات  $\sigma_1^2, \sigma_2^2, \dots, \sigma_n^2$  على الترتيب ، فعندئذ يكون للمتغير العشوائى :

$$Y = a_1 X_1 + a_2 X_2 + \dots + a_n X_n$$

توزيعاً طبيعياً بالتوقع :

$$\mu_y = a_1 \mu_1 + a_2 \mu_2 + \dots + a_n \mu_n$$

والتباين :

$$\sigma_y^2 = a_1^2 \sigma_1^2 + \dots + a_n^2 \sigma_n^2$$

وهكذا يبدو واضحاً الآن أن التوزيعين البواسونى ، والطبيعى يتمتعان بالخاصة التالية :

مجموع أى عدد من هذه المتغيرات المستقلة هو من جديد متغير عشوائى له نفس شكل التوزيع الذى تعود إليه هذه المتغيرات . وتنطبق هذه الخاصة أيضاً على توزيع كاي - مربع .

### نظرية (٥،١٢)

إذا كان للمتغيرات المستقلة  $X_1, \dots, X_n$  توزيعاً من نوع توزيع كاي - مربع بعدد من درجات الحرية  $\nu_1, \dots, \nu_n$  على الترتيب . فعندئذ يكون للمتغير

$$Y = X_1 + \dots + X_n$$

توزيعاً من نوع كاي - مربع بعدد من درجات الحرية مساوٍ لـ  $\nu = \nu_1 + \dots + \nu_n$

## البرهان

بحسب النظرية (٥,١٠) نجد أن :

$$M_y(t) = M_{x_1}(t) \dots M_{x_n}(t)$$

وحسب التمرين (٥,٨) لدينا :

$$M_{x_i}(t) = \frac{1}{(1-2t)^{\frac{x_i}{2_i}}} : i = 1, 2, \dots, n$$

وبالتعويض نجد أن :

$$M_y(t) = \frac{1}{(1-2t)^{\frac{v_1 + \dots + v_n}{2}}}$$

## نتيجة (٥,١)

إذا كانت المتغيرات المستقلة  $X_1, \dots, X_n$  طبيعية بالتوقع  $\mu$  ، والتباين  $\sigma^2$  فنعدّد  
يكون للمتغير

$$Y = \sum_{i=1}^n \left( \frac{X_i - \mu}{\sigma} \right)^2$$

توزيما من نوع كاي - مربع بـ  $n$  = درجة من الحرية .

إن النتيجة السابقة تنتج مباشرة من التمرين (٥,٥) والنظرية (٥,١٢) .

## (٣,٥) العينة العشوائية Random sampling

يمكن أن نعبر عن نتيجة تجربة إحصائية إما بقيمة عددية أو بتمثيل وصفي . فمثلا عند إلقاء حجرى نرد ، إذا كنا نهم بمجموع العددين اللذين ظهرا فإن نتيجة التجربة تمثلها بقيمة عددية . ومع ذلك إذا أجريت فحوص دم لطلاب مدرسة معينة ، وإذا كنا

نهم بنوع الدم عند كل طالب ، فعندئذ يكون التمثيل الوصفي أكثر فائدة في هذه الحالة .  
فدم الإنسان يمكن تصنيفه بثمانية طرق . وهذه الطرق هي :  $A, B, O$  أو  $AB$  مع إشارة  
زائدة أو ناقص حسب وجود أو فقدان المولد المضاد  $Rh$  .

ويتعامل الإحصائيون قبل كل شيء مع الملاحظات العددية . ففي تجربة اختبار  
نوع الدم فإن الإحصائي يرمز للأعداد من 1 إلى 8 لنوع دم كل طالب مفحوص ، وبعد  
ذلك يسجل الأعداد المخصصة من أجل كل طالب ، ويجد عدداً من الملاحظات مساوياً  
لعدد الطلاب الموجودين في المدرسة . وبهذا يكون لدى الإحصائي عدداً متنبهاً من  
النتائج .

وفي تجربة إلقاء حجرى نرد إذا كان اهتمام الإحصائي ينصب على مجموع الوجهين  
الذين ظهرا ، وإذا أُلقي هذين الحجرين عدداً لانهائياً من المرات فسيحصل على مجموعة  
غير متنبية من العناصر ، وكل ملاحظة ستمثل نتيجة إلقاء معين .

إن المجموع الكلى للملاحظات التى تهمننا ، سواء كان متنبهاً أم غير متنبه يؤلف ما  
يدعى بالمجتمع (population) . في السنوات الماضية كانت كلمة مجتمع تعنى جماعة من  
الناس ، أما الآن فيستخدم الإحصائي هذا المصطلح لدراسة أى شيء يهتم به سواء كان  
جماعة من الناس أو من الحيوانات أو الأشياء .

### تعريف (٥,٢) المجتمع

إن المجموع الكلى للملاحظات التى نهتم بها يدعى بالمجتمع وعدد الملاحظات في  
مجتمع ما يُعرف لنا بما يسمى حجم المجتمع . فمثلاً إذا كان المجتمع المدروس هو إنتاج  
مصنع للأدوات الكهربائية ، وإذا كان الإنتاج العام لهذا المصنع في اليوم<sup>(١٠)</sup> لبة ،  
فعندئذ نقول بأن حجم المجتمع المدروس هو<sup>(١٠)</sup> . كذلك فإن عدد أوراق اللعب في  
ورق اللعب هو مجتمع متنبه . وعلى هذا الأساس فإن هناك نوعين من المجتمعات مجتمعاً  
متنبهاً ومجتمع غير متنبه .

إن كل ملاحظة في المجتمع تمثل قيمة لمتغير عشوائى  $X$  له نفس توزيع المجتمع  
 $f(x)$  ، وعندما نتكلم عن مجتمع طبيعي ، أو حداني ، أو بشكل عام مجتمع  $f(x)$  فإن

المقصود بذلك هو مجتمع ملاحظاته تمثل قيماً لمتغير عشوائى له التوزيع الطبيعي أو الحدائى أو التوزيع الاحتمالى  $f(x)$  . كذلك فإن توقع وتباين متغير عشوائى فى مجتمع  $f(x)$  هما نفس توقع وتباين المجتمع الموافق  $f(x)$  .

والذى يهم الإحصائى هو البحث بطريق الاستنتاج عن الوسطاء المجهولة للمجتمع . فمثلاً يحوى المجتمع الطبيعي متغيرين هما  $\mu$  و  $\sigma$  وقد يكون أحد هذين الوسيطين أو كلاهما مجهولاً . ولذلك فإن غاية الإحصائى تقدير هذين الوسيطين من المعلومات التى تقدمها العينة المستخرجة من المجتمع بصورة عشوائية . وهذا ما يقودنا إلى نظرية العينات . وإذا كانت استنتاجاتنا مضبوطة ( دقيقة ) ، فيجب علينا أن نفهم العلاقة بين العينة ومجتمعها ، وعلى العينة أن تكون ممثلة للمجتمع المأخوذة منه ، وعليها أن تكون عينة عشوائية ، بمعنى أن الملاحظات التى تتألف منها يجب أن تكون مستقلة وعشوائية .

لنأخذ مثلاً عينة حجمها  $n$  من المجتمع المدروس  $f(x)$  ، ولنعرف المتغير العشوائى .  
 $X_i, i = 1, 2, \dots, n$  الممثل للقياس  $i$  أو قيمة الملاحظة المدروسة . عندئذ تؤلف المتغيرات العشوائية  $X_1, X_2, \dots, X_n$  عينة عشوائية فى المجتمع  $f(x)$  بالقيم العددية  $x_1, \dots, x_n$  فيما إذا كانت القياسات مجردة فى تجربة مكررة  $n$  مرة تحت نفس الظروف ، وعلى هذا الأساس يمكن أن نفرض أن المتغيرات  $X_1, \dots, X_n$  مستقلة ، ولكل واحد منها نفس توزيع المجتمع  $f(x)$  ، وهذا يعنى أن توزيعات هذه المتغيرات هى  $f(x_1), \dots, f(x_n)$  على الترتيب وتوزيعها المشترك هو :

$$f(x_1, \dots, x_n) = f(x_1) \dots f(x_n)$$

#### تعريف (٥،٣) حجم عينة عشوائية Volume of random sample

إذا كان للمتغيرات المستقلة  $X_1, \dots, X_n$  نفس التوزيع الاحتمالى  $f(x)$  ، عندئذ نقول بأن المتغيرات  $X_1, \dots, X_n$  تؤلف عينة عشوائية حجمها  $n$  من المجتمع  $f(x)$  ، ونكتب توزيعها المشترك على النحو التالى .

$$f(x_1, \dots, x_n) = f(x_1) \dots f(x_n)$$

### ٥,٤) نظرية العينات Sampling theory

إن هدفنا الرئيسى من سحب عينة عشوائية هو الحصول على معلومات حول الوسطاء المجهولة فى المجتمع المدروس . لنفرض أننا نريد الوصول عن طريق الاستنتاج لمعرفة نسبة الناس الذين يفضلون نوعا معينا من القهوة فى مدينة معينة . فمن المستحيل أن نسأل كل إنسان فى هذه المدينة عن رأيه وإحصاء المتغير الممثل للنسبة الحقيقية للأشخاص الذين يفضلون هذا النوع من القهوة . وعوضا عن ذلك ، فإننا نأخذ عينة عشوائية كبيرة من الناس ونسأل كل واحد منهم عن رأيه ، ونستنتج نسبة الذين يفضلون هذا النوع من القهوة . إن هذه القيمة سنستخدمها لاستنتاج بعض الحقائق المتعلقة بالنسبة الحقيقية فى هذه المدينة .

إن القيمة المحسوبة من خلال عينة تدعى إحصاء . وبما أنه يمكن أن نستخرج من مجتمع واحد عدة عينات عشوائية وبالتالي عدة إحصاءات ، لذلك علينا أن نتوقع بعض الاختلاف البسيط فى هذه الإحصاءات من عينة إلى أخرى . ولذلك فإن الإحصاء هو متغير عشوائى .

#### تعريف (٥,٤)

الإحصاء هو متغير عشوائى يعتمد فقط على قيم العينة العشوائية . نرسم عادة للإحصاء بأحد الأحرف اللاتينية الكبيرة . إن نسبة العينة فى المثال السابق هو إحصاء نرسم له عادة بالحرف  $P$  . كما أن قيمة المتغير العشوائى  $P$  ( من أجل عينة ما ) والتي نرسم لها بالرمز  $P$  تستخدم كتقدير للنسبة الحقيقية  $P$  من الناس فى المدينة المفروضة والذين يفضلون هذا النوع من القهوة ، وعلينا أن نعرف التوزيع الاحتمال للإحصاء  $P$  .

سنعرف إحصاءين هامين هما الوسط (the mean) والمتوسط (the median) . ويعتبر الوسط من أهم الإحصاءات لأنه يعبر عن تمرکز التوزيع ، وهو أحد الإحصاءات المفيدة والأكثر استخداما ، وهو ما يشار إليه باسم الوسط الحسابى .

**تعريف (٥,٥) وسط العينة Sample mean**

لتكن  $X_1, \dots, X_n$  عينة عشوائية حجمها  $n$  . نُعرف وسط العينة (sample mean) بالإحصاء كما يلي :

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$$

نلاحظ أن الإحصاء  $\bar{X}$  يفترض القيمة  $\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$  ( عندما يفترض  $X_1$  القيمة  $x_1$  ،  $X_2$  القيمة  $x_2$  ، وهلم جرا ) .

**مثال (٥,٩)**

أوجد وسيط العينة الممثلة للملاحظات 15, 17, 19 .

**الحل**

إن القيمة الملاحظة للإحصاء  $\bar{X}$  هي :

$$\bar{X} = \frac{15 + 17 + 19}{3} = 17$$

ونذكر هنا أن الإحصاء  $\bar{X}$  سيستخدم كتقدير لمتوسط المجتمع  $\mu$  المجهول . والإحصاء المفيد الثاني ، والذي يقيس لنا مركز مجموعة من البيانات ، يدعى بالوسيط ، وسنرمز له بالرمز  $\bar{X}$  .

**تعريف (٥,٦) متوسط العينة Sample median**

لتكن  $X_1, \dots, X_n$  عينة عشوائية حجمها  $n$  مرتبة تصاعدياً وفقاً لقيمتها . نُعرف متوسط العينة (sample median) بأنه الإحصاء :



$$X = \begin{cases} \frac{\frac{X}{2} + X \left( \frac{n}{2} \right) + 1}{2} & : \text{إذا كانت } n \text{ زوجية} \\ X \frac{n+1}{2} & : \text{إذا كانت } n \text{ فردية} \end{cases}$$

مثال (٥,١٠)

أوجد متوسط العينة العشوائية 7, 3, 4, 5, 5, 8, 7

الحل

بترتيب الملاحظات ترتيباً تصاعدياً وفقاً لقيمتها كالاتي : 3, 4, 5, 5, 7, 7, 8 .  
والملاحظ أن  $n = 7$  فردية لذلك فإن  $\tilde{x} = \frac{x_{n+1}}{2}$  ومنه  $\tilde{x} = x_4 = 5$  إذن فالوسط هو  $\tilde{x} = 5$

مثال (٥,١١)

أوجد متوسط عينة عشوائية ملاحظات 9, 4, 5, 6

الحل

لنرتب قبل كل شيء هذه الملاحظات ترتيباً تصاعدياً وفقاً لقيمتها فنجدها كالاتي 4, 5, 6, 9 .  
ونلاحظ أن  $n = 4$  زوجية إذا

$$\tilde{x} = \frac{\frac{x_4}{2} + \frac{x_5}{2} + 1}{2} = \frac{x_2 + x_3}{2}$$

ومنه  $\tilde{x} = \frac{6+5}{2} = 5.5$  وأخيراً فإن الوسيط هو  $\tilde{x} = 5.5$

بعد أن حددنا تمركز توزيع المعلومات الإحصائية لنحاول الآن أن نقدم قياساً لتباين هذه المعلومات . لنأخذ مثلاً مجموعة الملاحظات الآتية :

3, 4, 5, 6, 8, 9, 10, 12, 15

3, 8, 8, 9, 9, 9, 10, 10, 15

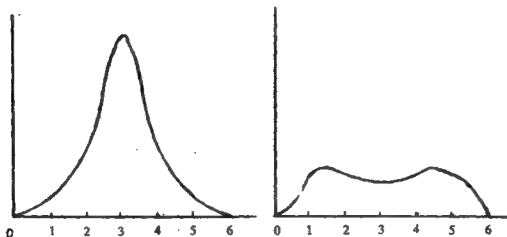
فلاحظ أن لهما نفس المدى 12 ( يعرف المدى بأنه الفرق بين أكبر وأصغر رقم ) . كما نلاحظ في مجموعة الملاحظات الأولى أن المتوسط والوسط متساويان وكلاهما يساوي 8 ، إلا أنه يوجد فرق كبير في تباعد القياسات عن الوسط . فبالنسبة للمجموعة الثانية من الملاحظات ، فإننا نلاحظ أن المتوسط والوسط متساويان ويساوي كلا منهما العدد 9 ، ولكن غالبية الأعداد قريبة من الوسط . ولخاصة التغير هذه أهمية كبيرة في المعلومات الإحصائية . فبالإضافة لأهمية العملية فهي تشكل قياساً ضرورياً إلى جانب قياس النزعة المركزية ( وهو القياس المعبر عن موضع تمركز التوزيع أى الوسط الحسابي مثلاً ) لبناء صورة ذهنية لتوزيع التواتر . هذا وتوجد قياسات عديدة للتباين سنناقش هنا أهمها ، وسنبداً بأبسطها وهو المدى .

**تعريف (٥,٧) المدى The range**

إن مدى عينة عشوائية  $X_1, \dots, X_n$  مرتبة تصاعدياً وفقاً لقيمتها يتحدد بالإحصاء  $X_n - X_1$  .

**مثال (٥,١٢)**

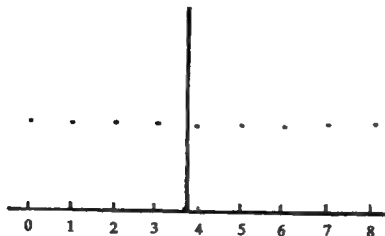
إن مدى الملاحظات 10, 12, 12, 18, 19, 22, 24 هو  $24 - 10 = 14$  . إذا كان حجم العينة كبيراً بقدر كاف ، فنعتقد بعتبر المدى قياساً رديئاً لتغيرات العينة ، لأنه لا يقدم لنا أى معلومات عن الطريقة التي تتوزع بها هذه القيم فيما بينها ، ولهذا السبب يعتبر المدى قياساً غير مرضٍ تماماً للتباين . لتأمل على سبيل المثال التوزيعين الموضحين على الشكل (٥,٤) .



الشكل (٥, ٤)

فلاحظ أن لكل منهما نفس المدى إلا أن تباين الشكل ٥, ٤ أكبر من تباين الشكل ٥, ٤ ب .  
 لنرى الآن فيما إذا كان بالإمكان إيجاد قياس للتباين يمكن التعبير عنه من خلال عدد ، وبحيث  
 يكون أكثر حساسية من المدى . لذلك نأخذ على سبيل المثال مجموعة القياسات 5, 7, 1, 2, 4 .  
 يمكن تمثيل هذه القياسات بيانياً كما في الشكل (٥, ٥) بوضع نقاط من أجل القياسات على طول  
 محور الفواصل . ولحساب الوسط نجد أن :

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} = \frac{19}{5} = 3.8$$



الشكل (٥, ٥)

وتمثل الوسط بنقطة على محور الفواصل . ويمكن النظر إلى التباين الآن من خلال المسافة بين كل النقاط ( القياسات ) والوسط  $\bar{X}$  . فإذا كانت المسافة كبيرة قلنا أن المعلومات الإحصائية أكثر تبايناً مما لو كانت المسافات أصغر . ونُعرّف انحراف قياس  $X_i$  بأنه المسافة بين هذا القياس والوسط  $(X_i - \bar{X})$  . ونلاحظ أن انحرافات القياسات الواقعة على يمين الوسط موجبة ، بينما انحرافات القياسات الواقعة على يسار الوسط سالبة . وفي المثال السابق يمكن تمثيل الانحرافات بالجدول (٥،١) .

$X_i$	$X_i - \bar{X}$	$(X_i - \bar{X})^2$	$X_i^2$
5	1.2	1.44	25
7	3.2	10.24	49
1	-2.8	7.84	1
2	-1.8	3.24	4
4	0.2	0.04	16
$\sum_{i=1}^5 X_i = 19$	0	22.8	95

فإذا اتفقنا على أن الانحرافات تقدم معلومات عن التباين ، فإن من واجبنا الآن أن نضع علاقة واضحة عن هذه الانحرافات تزودنا بمقياس جيد للتباين .

وكخطوة جيدة في هذا الطريق نختار ما يسمى بمربع الانحرافات .

**تعريف (٥،٨) تباين العينة Sample variance**

لتكن  $X_1, \dots, X_n$  عينة عشوائية حجمها  $n$  عندئذ نعرف تباين العينة بالإحصاء :

$$S^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n - 1}$$

لنرمز لقيمة  $S^2$  من أجل عينة معينة بالرمز  $S^2$  . والملاحظ أن  $S^2$  تمثل وسط مربعات انحرافات الملاحظات عن أوسطها . أما سبب استخدام  $(n - 1)$  عوضاً عن  $n$  ، فيظهر في الفصل السادس .

### نظرية (٥, ١٣)

إذا كان  $S^2$  ممثلاً لتباين عينة عشوائية ذات حجم  $n$  فعندئذ يمكننا أن نكتب  $S^2$  على النحو التالي :

$$S^2 = \frac{n \sum_{i=1}^n X_i^2 - \left( \sum_{i=1}^n X_i \right)^2}{n(n - 1)}$$

### البرهان

من تعريف  $S^2$  نستطيع أن نكتب :

$$\begin{aligned} S^2 &= \frac{\sum_{i=1}^n (X_i^2 - 2X_i\bar{X} + \bar{X}^2)}{n-1} \\ &= \frac{\sum_{i=1}^n X_i^2 - 2\bar{X} \sum_{i=1}^n X_i + n\bar{X}^2}{n-1} \end{aligned}$$

وبتبديل  $\bar{X} = \frac{\sum X_i}{n}$  ، وضرب كل من بسط ومقام الكسر السابق بالعدد  $n$  فإننا نجد أن :

$$S^2 = \frac{n \sum_{i=1}^n X_i^2 - (\sum X_i)^2}{n(n - 1)}$$

لنرمز للانحراف المعياري العيني بالرمز  $S$  ، عندئذ يمثل  $S$  الجذر الموجب للتباين العيني .

### مثال (٥, ١٣)

بين أن التباين العيني للملاحظات 6, 5, 7, 8, 9, 6 هو  $S^2 = 65$

الحل

نلاحظ أن :

$$\sum_{i=1}^6 X_i^2 = 291 \quad , \quad \sum_{i=1}^6 X_i = 41 \quad , \quad n=6$$

وحسب النظرية (٥, ١٣) نجد أن :

$$S^2 = \frac{(6)(291) - (41)^2}{(6)(5)} = 65$$

مثال (٥, ١٤)

احسب  $\bar{X}$ ,  $S$  لمجموعة القياسات 85, 90, 60, 70, 81

$X_i$	$X_i^2$
85	7225
70	4900
60	3600
90	8100
81	6561
386	30386

الحل

نلاحظ أن  $\bar{X} = \frac{386}{5} = 77.5$  كما نلاحظ أن :

$$\sum_{i=1}^n (X_i - \bar{X})^2 = \sum X_i^2 - \frac{(\sum X_i)^2}{n}$$

ولذلك فإن :

$$s = \sqrt{\frac{\sum (X_i - \bar{X})^2}{n-1}} = \sqrt{\frac{586.8}{4}} = \sqrt{146.7} = 12.1$$

يشكل حقل الإحصاء الاستقرائي قاعدة تتعلق بما يسمى بالتعميم والتنبؤ ، هذا ويمكن إجراء التعميم من الإحصاء إلى المتغير بثقة . إذا فهمنا تغير سلوك هذا الإحصاء عند حسابه من خلال عينات مختلفة مأخوذة من نفس المجتمع المدروس . وسيعتمد توزيع الإحصاء على حجم المجتمع ، وحجم العينة ، والطريقة التي تمت بها اختيار العينة الإحصائية . فإذا كان حجم المجتمع المدروس كبيراً جداً أو غير منته ، فعندئذ سيكون للإحصاء نفس التوزيع سواء سحبنا عينتنا مع الإعادة أو بدون إعادة . أما إذا احتوى مجتمعنا المدروس على عدد بسيط منته من العناصر ، فإن السحب مع الإعادة من هذا المجتمع سيؤدي إلى توزيع يختلف بشكل تافه عما إذا كنا سحبنا بدون إعادة . أن السحب مع الإعادة من مجتمع منته يكافئ سحباً من مجتمع غير منته . لذلك فإنه ليس هناك تحديد للحجم الممكن للعينة المسحوبة .

**تعريف (٥,١٠) التوزيع الاحتمالي Probabilistics distribution**

إن التوزيع الاحتمالي للإحصاء يسمى بتوزيع المعاينة .

**تعريف (٥,١١) الانحراف المعياري Standard deviation**

يدعى الانحراف المعياري لتوزيع المعاينة لإحصاء ما بالخطأ المعياري لهذا الإحصاء .

### ملاحظة

إن التوزيع الاحتمالي للإحصاء  $\bar{X}$  يسمى بتوزيع المعاينة للوسط الحسابي ، والخطأ المعياري للوسط هو الانحراف المعياري لتوزيع المعاينة لـ  $\bar{X}$  . إن كل عينة ذات حجم  $n$  مأخوذة من مجتمع معين ستقودنا إلى قيمة  $s$  للإحصاء  $S$  ، وتسمى هذه القيمة بالانحراف المعياري لتوزيع المعاينة . وعلى هذا الأساس فإن الخطأ المعياري للانحراف المعياري العيني ما هو إلا الانحراف المعياري للإحصاء  $S$  .

فيما تبقى من هذا الفصل سنورد بعض توزيعات المعاينة الهامة والتي يستخدمها الإحصائي بشكل تكرر. أما تطبيقات هذه التوزيعات ، فنسورها في الفصلين السادس والسابع .

### (٥,٥) توزيع المعاينة للوسط

#### Sampling distribution of mean

إن أول وأهم توزيعات المعاينة هو توزيع الوسط  $\bar{X}$  . لنفرض ، على سبيل المثال ، عينة عشوائية ذات حجم  $n$  مأخوذة من مجتمع طبيعي بالمتوسط  $\mu$  والتباين  $\sigma^2$  . إن كل ملاحظة  $X_i$  ,  $i = 1, n$  من العينة العشوائية سيكون لها نفس التوزيع الطبيعي . لذلك حسب تعريف الوسط :

$$\bar{X} = \frac{X_1 + \dots + X_n}{n}$$

وحسب النظرية (٥,١١) نجد أن  $\bar{X}$  توزيعاً طبيعياً بالتوقع :

$$\bar{\mu} = \frac{\mu_1 + \mu_2 + \dots + \mu_n}{n} = \mu$$

والتباين :

$$\sigma_{\bar{x}}^2 = \frac{\sigma_1^2 + \sigma_2^2 + \dots + \sigma_n^2}{n^2} = \frac{\sigma^2}{n}$$

إذا سحبنا عينة من مجتمع توزيعه مجهول متته أو غير متته ، فمعتقد سيكون توزيع المعاينة للإحصاء  $\bar{X}$  قريباً جداً من التوزيع الطبيعي بالتوقع  $\mu$  والتباين  $\frac{\sigma^2}{n}$  شريطة أن يكون حجم العينة  $n$  كبيراً . إن هذه النتيجة المذهلة تنتج مباشرة من النظرية التالية والتي نسميها عادة بنظرية النهايات المركزية (central limit theorem) والتي نؤجل برهانها في هذا الكتاب .



## نظرية (٥,١٤)

ليكن  $\bar{X}$  وسط عينة عشوائية حجمها  $n$  مأخوذة من مجتمع إحصائي بالتوقع  $\mu$  والتباين  $\sigma^2$ . عندئذ يسمى التوزيع الاحتمالي للمتغير :

$$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$$

إلى التوزيع الطبيعي المعياري عندما تسعى  $n$  إلى  $\infty$ .

إن التقريب الطبيعي للإحصاء  $\bar{X}$  سيكون جيداً إذا كان  $n \geq 30$  بغض النظر عن شكل المجتمع. أما إذا كان  $n < 30$  فعندئذ سيكون التقريب جيداً فيما إذا كان المجتمع المدروس لا يختلف كثيراً عن المجتمع الطبيعي. أما إذا علم أن المجتمع طبيعياً، فعندئذ سيكون للإحصاء توزيع معانية طبيعي مهما كان حجم العينة  $n$ .

## مثال (٥,١٥)

تنتج شركة للمصابيح الكهربائية نوعاً من اللامبات عمرها (مدة عملها) له تقريباً توزيعاً طبيعياً بالوسط  $\mu = 750$  ساعة، والانحراف المعياري  $\sigma = 38$  ساعة. أوجد احتمال أن يكون لعينة مؤلفة من 14 لمبة وسطاً أقل من 735 ساعة.

## الحل

نلاحظ أن توزيع المعاينة لوسط العينة  $X$  سيكون قريباً من الطبيعي بـ  $\mu_x = 750$ ،  $\sigma_x = \frac{38}{\sqrt{14}} = 10.155$ . كما نلاحظ أن الاحتمال المطلوب يمثل مساحة القسم المظلل على الشكل (٥,٦).

واعتماداً على كون  $\bar{X}_1 = 735$  فإننا نجد أن :

$$Z_1 = \frac{735 - 750}{10.155} = -1.476$$

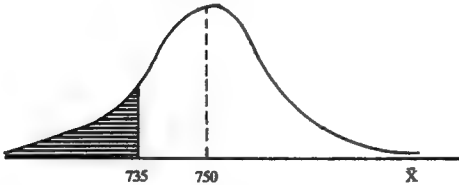
ولذلك فإن :

$$P[\bar{X} < 735] = P[Z < -1.48]$$

وحسب الجدول IV نجد أن :

$$P[\bar{X} < 735] = 0.0694$$

$$= 0.0694$$



الشكل (٥,٦)

مثال (٥,٦)

بفرض أن المجتمع الإحصائي المدروس هو المجتمع المنقطع المنتظم بالكثافة :

$$f(x) = \begin{cases} \frac{1}{6} , & : x = 1, 2, 3, 4, 5, 6 \\ 0 , & : \text{فيما عدا ذلك} \end{cases}$$

أوجد احتمال أن نسحب عينة حجمها  $n$  من هذا المجتمع مع الإعادة . فنجد أن وسط العينة أقل من 3.7 وأكبر من 3.2 إذا قيس الوسط بالنسبة لأقرب جزء عشري .

الحل

لحساب وسط وتباين التوزيع المنتظم السابق فإننا نجد أن :

$$\mu_x = \frac{1 + 2 + 3 + 4 + 5 + 6}{6} = 3.5$$

$$\sigma_x^2 = \frac{(1 - 3.5)^2 + (2 - 3.5)^2 + \dots + (6 - 3.5)^2}{6} = 2.91$$

هذا ويمكن تقريب توزيع المعاينة لـ  $\bar{X}$  إلى التوزيع الطبيعي بالوسط  $\mu = 3.5$  والتباين  $\sigma^2 = 0.485$  ، وبأخذ الجذر التربيعي نجد أن  $\sigma = 0.696$  .

أما احتمال أن يكون  $\bar{X}$  أقل من 3.7 وأكبر من 3.2 فيمثل بالقسم المظلل من الشكل (٥،٧) ، وقيم  $z$  الموافقة لـ  $\bar{X}_1 = 3.15$  و  $\bar{X}_2 = 3.75$  فهي على الترتيب .

$$Z_1 = \frac{3.15 - 3.5}{0.696} = -0.502$$

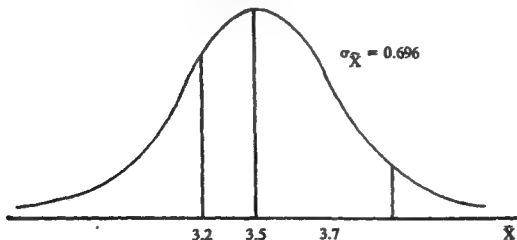
$$Z_2 = \frac{3.75 - 3.5}{0.696} = 0.359$$

لذلك فإن :

$$P[3.2 < \bar{X} < 3.7] \simeq P[-0.502 < Z < 0.359]$$

$$= P[Z < 0.359] - P[Z < -0.502]$$

$$\simeq 0.6406 - 0.3085 = 0.3321$$



الشكل (٥،٧)

نفرض أن لدينا مجتمعين ، الأول بالوسط  $\mu_1$  والتباين  $\sigma_1^2$  أما الثاني بـ  $\bar{X}_1$  فالوسط  $\mu_2$  والتباين  $\sigma_2^2$  . لنرمز بـ  $\bar{X}_2$  لوسط عينة حجمها  $n_1$  مسحوبة بشكل عشوائى من المجتمع الأول وبـ  $\bar{X}_2$  لوسط عينة حجمها  $n_2$  مسحوبة من المجتمع الثاني وبشكل عشوائى أيضا . وبصورة مستقلة عن العينة الأولى . والسؤال المطروح الآن : كيف يتوزع الفرق  $\bar{X}_2 - \bar{X}_1$  عند تكرار السحب ؟

للإجابة على هذا السؤال نلاحظ حسب النظرية (٥,١٤) أن للمتغيرين  $\bar{X}_1$  ،  $\bar{X}_2$  توزيعين قريين من التوزيع الطبيعي بالوسطين  $\mu_1$  ،  $\mu_2$  ، والتباينين  $\frac{\sigma_1^2}{n_1}$  ،  $\frac{\sigma_2^2}{n_2}$  على الترتيب . ويتحسن هذا التقريب بازدياد كل من  $n_1$  ،  $n_2$  وباختيار عينات مستقلة من المجتمعين نجد أن المتغيرين :  $\bar{X}_2$  ،  $\bar{X}_1$  سيكونان مستقلين وعندئذ بحسب النظرية (٥,١١) وبفرض أن  $q_1 = 1$  و  $q_2 = -1$  نستنتج أن للمتغير  $\bar{X}_1 - \bar{X}_2$  توزيعاً قريباً من التوزيع الطبيعي بالوسط :

$$\mu_{\bar{X}_1 - \bar{X}_2} = \mu_{\bar{X}_1} - \mu_{\bar{X}_2} = \mu_1 - \mu_2$$

والتباين :

$$\sigma_{\bar{X}_1 - \bar{X}_2}^2 = \sigma_{\bar{X}_1}^2 + \sigma_{\bar{X}_2}^2 = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$$

### نظرية (٥,١٥)

لدى سحب عينات مستقلة حجمها  $n_1$  ،  $n_2$  من مجتمعين منقطعين أو مستمرين بالوسطين  $\mu_1$  ،  $\mu_2$  والتباينين  $\sigma_1^2$  ،  $\sigma_2^2$  على الترتيب ، فعندئذ سيكون للفرق  $\bar{X}_1 - \bar{X}_2$  توزيعاً قريباً من التوزيع الطبيعي بالوسط  $\mu_1 - \mu_2$  والتباين  $\sigma_{\bar{X}_1 - \bar{X}_2}^2$  :

$$\sigma_{\bar{X}_1 - \bar{X}_2}^2 = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$$

وحسب نظرية النهايات المركزية سيكون للمتغير :

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

توزيعاً قريباً من التوزيع الطبيعي المعياري .

إذا كان كلا من  $n_1$  ,  $n_2$  أكبر من 30 ، فعندئذ سيكون التقريب الطبيعي لتوزيع  $\bar{X}_1 - \bar{X}_2$  جيداً جداً .

مثال (٥،١٧)

بفرض أن لقنوات التصوير التلفزيونية العائدة للمنتج A وسط عمر قدره ست سنوات ونصف وانحراف معياري قدره 0.9 سنة . بينما لها وسط عمر قدره ست سنوات وانحراف معياري قدره 0.8 سنة بالنسبة لمصنع يعود إلى المنتج B . ما هو احتمال أن يكون لعينة عشوائية مؤلفة من 36 قناة ( عائدة للمنتج A ) على الأقل وسط عمر أكبر بعام من وسط عمر عينة عشوائية مؤلفة من 49 قناة عائدة للمنتج B ؟

الحل

لنرتب بعض المعلومات السابقة في الجدول التالي :

المجتمع الأول	المجتمع الثاني
$\mu_1 = 6.5$	$\mu_2 = 6.0$
$\sigma_1 = 0.9$	$\sigma_2 = 0.8$
$n_1 = 36$	$n_2 = 49$

باستخدام النظرية (٥،١٥) نلاحظ أن لتوزيع المعاينة للفرق  $\bar{X}_1 - \bar{X}_2$  وسطاً وانحرافاً معيارياً قدرهما :

$$\sigma_{\bar{x}_1 - \bar{x}_2} = 6.5 - 6.0 = 0.5$$

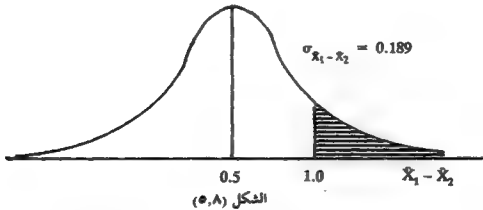
$$\sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{0.81}{36} + \frac{0.64}{49}} = 0.189$$

نلاحظ أن احتمال أن يكون لـ 36 قناة العائدة للمنتج A وسطا على الأقل أكبر بعام من وسط الـ 49 قناة العائدة للمنتج B يمكن تمثيله بالقسم المظلل على الشكل الموضح (٥, ٨) . ويوافق القيمة  $\bar{x}_1 - \bar{x}_2 = 1.0$  :

$$Z = \frac{1.0 - 0.5}{0.189} = 2.646$$

لذلك فإن :

$$\begin{aligned} P[\bar{x}_1 - \bar{x}_2 \geq 1.0] &= P[Z > 2.646] \\ &= 1 - P[Z < 2.646] \\ &= 1 - 0.9959 \\ &= 0.0041 \end{aligned}$$



### (٥, ٦) توزيع المعاينة للمغير العشوائي $\frac{(n-1)S^2}{\sigma^2}$

إذا سحبنا عينة عشوائية ذات حجم  $n$  من مجتمع طبيعي بالوسط  $\mu$  والتباين  $\sigma^2$  ، وحسبنا تباين العينة  $S^2$  ، فإننا نحصل على قيمة للإحصاء  $S^2$  . وللإحصاء  $S^2$  تطبيقات

عملية قليلة . وعوضاً عن هذا الإحصاء سندرس توزيع المتغير العشوائى  $\frac{(n-1) S^2}{\sigma^2}$

بجمع وطرح العينة  $\bar{X}$  العبارة  $\sum_{i=1}^n (X_i - \mu)^2$  فإننا نجد أن :

$$\begin{aligned} \sum_{i=1}^n (X_i - \mu)^2 &= \sum_{i=1}^n [(\bar{X}_i - \bar{X}) + (\bar{X} - \mu)]^2 \\ &= \sum_{i=1}^n (X_i - \bar{X})^2 + \sum_{i=1}^n (\bar{X} - \mu)^2 + 2(\bar{X} - \mu) \cdot \sum_{i=1}^n (X_i - \bar{X}) \\ &= \sum_{i=1}^n (X_i - \bar{X})^2 + n(\bar{X} - \mu)^2 \end{aligned}$$

وبتقسيم طرفى العلاقة السابقة على  $\sigma^2$  وبوضع  $(n-1)S^2$  بدلاً عن  $\sum_{i=1}^n (X_i - \bar{X})^2$  فإننا نجد أن :

$$\frac{\sum_{i=1}^n (X_i - \mu)^2}{\sigma^2} = \frac{(n-1) S^2}{\sigma^2} + \frac{(\bar{X} - \mu)^2}{\sigma^2/n}$$

وحسب النتيجة (٥,١) نعلم أن للمتغير  $\frac{\sum_{i=1}^n (X_i - \mu)^2}{\sigma^2}$  توزيعاً من نوع كاي - مربع بـ  $n$  درجة من الحرية . أما المتغير  $\frac{(\bar{X} - \mu)^2}{\sigma^2/n}$  فله توزيع كاي - مربع بدرجة حرية واحدة وذلك حسب الثمرين (٥,٥) وبما أن فرق متغيرين مستقلين لهما توزيع كاي - مربع هو من جديد متغير من نوع كاي مربع وبذلك يكون للمتغير  $\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{\sigma^2}$  توزيعاً من نوع كاي - مربع بعدد درجات الحرية مساو لـ  $(n-1)$  .

### نظرية (٥,١٦)

إذا كان  $S^2$  ممثلاً لتباين عينة عشوائية حجمها  $n$  مأخوذة من مجتمع طبيعى بالتباين  $\sigma^2$  ، فعندئذ سيكون للمتغير العشوائى :

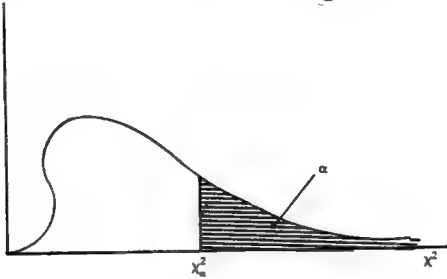
$$\chi^2 = \frac{(n-1) S^2}{\sigma^2}$$

توزيعاً من نوع كاي - مربع بـ  $p = n - 1$  درجة من الحرية .

إن قيم المتغير العشوائى  $\chi$  تحسب بالنسبة لكل عينة من العلاقة :

$$\chi^2 = \frac{(n-1) S^2}{\sigma^2}$$

كما أن احتمال أن تقدم عينة عشوائية قيمة  $\chi^2$  أكبر من قيمة مفروضة يساوى إلى المساحة الواقعة تحت المنحنى والواقعة إلى يمين المفروضة ، ومن المعتاد أن نرمز بـ  $\chi^2_{\alpha}$  لقيمة  $\chi^2$  التى تحدد لنا مساحة على يمينها قدرها  $\alpha$  . وهذه المساحة يوضحها القسم المظلل على الشكل (٥,٩) الموضح أدناه .



الشكل (٥,٩)

الجدول السادس يوضح القيم المختلفة لـ  $\chi^2_{\alpha}$  الموافقة للقيم المختلفة لكل من  $\nu$  ،  $\alpha$  . ونلاحظ في هذا الجدول أن العمود الأيسر يوافق قيم  $\nu$  المختلفة من 1 إلى 30 ، أما السطر العلوى فيعطينا المساحات  $\alpha$  . لذلك فإن قيمة  $\chi^2$  ( من أجل  $\nu = 10$  ) والتي تحدد مساحة على يمينها قدرها 0.01 هي  $\chi^2_{0.01} = 23.209$  . وبسبب عدم وجود تناظر في الشكل ، علينا أن نبحث أيضا عن  $\chi^2_{0.99} = 2.558$  من أجل  $\nu = 10$  .

أن 99% من توزيع كاي - مربع بـ  $n - 1$  درجة من الحرية يقع بين النقطتين  $\chi^2_{0.995}$  و  $\chi^2_{0.005}$  ويشير إلى أن قيمة  $\chi^2$  الواقعة على يمين  $\chi^2_{0.995}$  ليست مرجحة للوقوع إلا إذا افترضنا قيمة لـ  $\chi^2$  صغيرة جدا . وبشكل مشابه فإن قيمة  $\chi^2$  الواقعة على يسار ليست مرجحة للوقوع إلا إذا افترضنا قيمة لـ  $\chi^2$  كبيرة جدا . وبعبارة أخرى



يمكن أن يكون لـ  $\chi^2$  قيمة مرافقة على يسار القيمة  $\chi^2_{0.995}$  أو على يمين القيمة  $\chi^2_{0.005}$  إذا كانت  $\sigma^2$  مضبوطة . وإذا صدف وحدث هذا فعلى الغالب نكون قد ارتكبنا خطأ في قيمة  $\sigma^2$  المفروضة .

### مثال (٥, ١٨)

تكفل شركة صناعية لإنتاج بطاريات السيارات بالعمل في المتوسط أربع سنوات وبالاختلاف المعياري سنة . فإذا جربت أربع بطاريات ووجد أن عمر كل منها على النحو التالي 4.0, 3.8, 2.8, 2.3 سنة ، فهل ستظل الشركة مقتنعة من أن لبطارياتها انحراف معياري قدره عام ؟

### الحل

لنبحث أولاً عن تباين العينة :

$$S^2 = \frac{(4)(43.57) - (12.9)^2}{(4)(3)} = 0.6558$$

كما أن :

$$\chi^2_1 = \frac{(3)(0.6558)}{1} = 1.9675$$

تمثل قيمة كاي - مربع بثلاث درجات من الحرية . لذلك فإن 99% من قيم  $\chi^2$  بثلاث درجات من الحرية ستقع بين النقطتين 0.0717, 12.838 والمحسوبتين من أجل  $\sigma = 1$  وهي معقولة ، لذلك لا يوجد أى سبب لدى الشركة المنتجة للشك في أن الانحراف المعياري يمثل قيمة أخرى غير العام .

### (٥, ٧) التوزيع t - Distribution

قد لانكون معظوظين في كثير من الأحيان في معرفة تباين المجتمع الذى نختار منه العينات العشوائية . فمن أجل عينات حجمها  $n \geq 30$  يكون  $S^2$  تقديراً جيداً لتباين

المجتمع  $\sigma^2$  والسؤال المطروح الآن : هل يحدث تغيراً في توزيع الإحصاء  $\frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$  في النظرية (٥, ١٤) إذا بدلنا  $\sigma^2$  بـ  $S^2$  ! بما أن  $S^2$  تزودنا بتقدير جيداً لـ  $\sigma^2$  ولا تتغير من عينة لأخرى من أجل  $n \geq 30$  فإن توزيع الإحصاء  $\frac{\bar{X} - \mu}{S / \sqrt{n}}$  يظل قريباً من التوزيع الطبيعي المعياري .

أما إذا كان حجم العينة المختارة ( $n < 30$ ) فإن قيم  $S^2$  متغير من عينة لأخرى وتوزيع المتغير العشوائي  $\frac{\bar{X} - \mu}{S / \sqrt{n}}$  لن يكون قريباً من التوزيع الطبيعي المعياري . سنذكر الآن متغيراً عشوائياً جديداً نرسم له بالرمز  $T$  ونعرفه على النحو التالي :

$$T = \frac{\bar{X} - \mu}{S / \sqrt{n}}$$

ولإيجاد توزيع المعاينة للمتغير  $T$  ، فإننا سنفترض أن العينة مسحوبة من مجتمع طبيعي ، وعندئذ نستطيع أن نكتب :

$$T = \frac{(\bar{X} - \mu) / (\sigma / \sqrt{n})}{\sqrt{S^2 / \sigma^2}} = \frac{Z}{\sqrt{v / (n - 1)}}$$

حيث يكون للمتغير :

$$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$$

توزيعاً طبيعياً معيارياً وللمتغير :

$$v = \frac{(n - 1) S^2}{\sigma^2}$$

توزيعاً من نوع كاي - مربع بـ  $v = n - 1$  درجة من الحرية . ومن أجل عينات طبيعية ، يمكن أن نبين أن  $S^2$  و  $\bar{X}$  يمثلان متغيرين مستقلين ، وهكذا فإن  $Z$  ،  $v$  يمثلان متغيرين مستقلين أيضاً .

نظرية (٥, ١٧)

ليكن  $Z$  متغيراً عشوائياً طبيعياً معيارياً ،  $v$  متغير من نوع كاي - مربع بـ درجة من الحرية . إذا كان المتغيران  $Z$  و  $v$  مستقلين ، فنحن نؤكد أن يكون توزيع المتغير :

$$T = \frac{Z}{\sqrt{V/\nu}}$$

معروفة بالعلاقة :

$$h(t) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right) \cdot \sqrt{\nu \cdot \pi}} \cdot \left(1 + \frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}} : -\infty < t < +\infty$$

وهذا ما يعرف باسم التوزيع  $t$  بـ درجة من الحرية .

### البرهان

بما أن المتغيرين العشوائيين  $Z$  و  $V$  مستقلان ، لذلك فإن التوزيع الاحتمالي المشترك لهما يعطى بالعلاقة :

$$f(Z, V) = \begin{cases} \frac{1}{\sqrt{2\pi}} e^{-\frac{Z^2}{2}} \cdot \frac{1}{2^{\nu/2} \Gamma\left(\frac{\nu}{2}\right)} e^{-\frac{1}{2}V} V^{\frac{\nu}{2}-1} & : -\infty < Z < +\infty \\ 0 & 0 < V < \infty \end{cases}$$

فيما عدا ذلك :

لنعرف متغيراً جديداً  $U = V$  . إن الحل العكسي للمعادلتين  $u = v$  و  $t = \frac{Z}{\sqrt{v/\nu}}$  هما  $u = v$  ،  $Z = \frac{t\sqrt{u}}{\sqrt{\nu}}$  ، ونحصل منهما على :

$$J = \begin{bmatrix} \frac{\sqrt{u}}{\sqrt{\nu}} & \frac{t}{2\sqrt{u \cdot \nu}} \\ 0 & 1 \end{bmatrix} = \frac{\sqrt{u}}{\sqrt{\nu}}$$

ويمثل التقابل السابق تقابلاً واحداً لواحد بين عناصر المجموعة :

$$\{(Z, V) | -\infty < Z < +\infty, 0 < V < \infty\}$$

وعناصر المجموعة  $\{(t, u) | -\infty < t < +\infty, 0 < u < \infty\}$  ، وباستخدام النظرية (٥،٤) نجد أن التوزيع الاحتمالي المشترك للمتغيرين  $u$  و  $t$  هو من الشكل :

$$g(t, u) = \begin{cases} \frac{1}{\sqrt{2\pi} 2^{v/2} \Gamma(\frac{v}{2})} e^{\frac{v}{2}-1} \cdot e^{-\left(\frac{v}{2}\right)\left(1+\frac{t^2}{v}\right)} \cdot \frac{\sqrt{u}}{\sqrt{v}} & : -\infty < t < +\infty \\ & 0 < u < \infty \\ 0 & \text{فيما عدا ذلك} \end{cases}$$

وبالمكاملة بالنسبة لـ  $u$  نجد أن التوزيع الاحتمالي للمتغير  $T$  هو من الشكل :

$$h(t) = \int_0^\infty g(t, u) du = \int_0^\infty \frac{1}{\sqrt{2\pi} 2^{v/2} \Gamma(\frac{v}{2})} u^{\frac{v}{2}-1} \cdot e^{-\frac{v}{2}\left(1+\frac{t^2}{v}\right)} \frac{\sqrt{u}}{\sqrt{v}} du$$

$$Z = \frac{u\left(1 + \frac{t^2}{v}\right)}{2} \quad \text{لنجرى الآن تغيراً في المتحول بالعلاقة}$$

$$dv = dz \left/ \left(1 + \frac{t^2}{v}\right)\right. \quad \text{وهكذا فإن} :$$

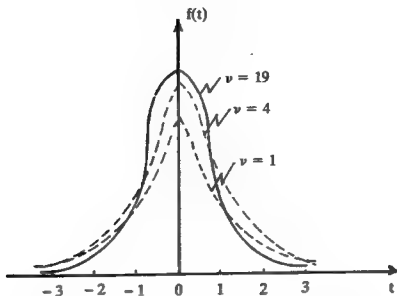
$$h(t) = \frac{1}{\Gamma\left(\frac{v}{2}\right) \sqrt{\pi v}} \left(1 + \frac{t^2}{v}\right)^{-\frac{v+1}{2}} \cdot \int_0^\infty Z^{\frac{v+1}{2}-1} \cdot e^{-Z} dz$$

$$= \frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right) \cdot \sqrt{\pi v}} \cdot \left(1 + \frac{t^2}{v}\right)^{-\frac{v+1}{2}} : -\infty < t < +\infty$$

فيما عدا ذلك :

$$= 0$$

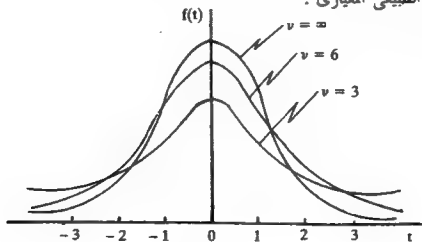
تدعى الدالة الأخيرة بدالة كثافة التوزيع  $t$  بـ  $v$  درجة من الحرية حيث يمثل عدداً صحيحاً موجباً . ويقدم الشكل (٥, ١٠) عدة تمثيلات بيانية لهذا المنحنى من أجل قيم متعددة لـ  $v$  .



الشكل (٥, ١٠)

نلاحظ أن التوزيع  $T$  متماثل بشكل مشابه للتوزيع  $Z$ . حيث أنهما متناظرين حول الوسط الذي يساوى الصفر. كما أن كلا التوزيعين له شكل المنحنى الجرسى، غير أن التوزيع  $T$  أكثر تغيراً، بسبب أن قيم  $T$  تعتمد على تغيرات قيم الكميتين  $\bar{X}$  و  $S^2$  في حين أن قيم  $Z$  تعتمد على تغير  $\bar{X}$  من عينة لأخرى. والملاحظ أيضاً أن توزيع  $T$  يختلف عن التوزيع  $Z$  في كون أن تباينه يعتمد على حجم العينة  $n$ ، وهو أكبر من الواحد دوماً. ويصبح التوزيعان متماثلين فقط في الحالة التي يسعى فيها  $n \rightarrow \infty$ .

ويوضح لنا الشكل (٥, ١١) درجة الصلة بين التوزيع  $t$  من أجل  $\nu = 3, 6$ ، والتوزيع الطبيعي المعياري.

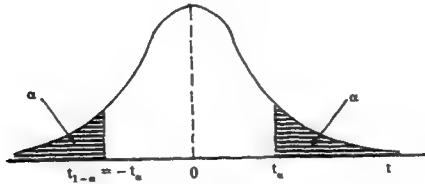


الشكل (٥, ١١)

إن احتمال أن تقدم العينة العشوائية قيمة  $t = (\bar{X} - \mu) / (S / \sqrt{n})$  واقعة بين قيمتين مختلفتين يساوى إلى المساحة الواقعة تحت منحنى التوزيع  $t$  بين الاحداثين السنين لهاتين القيمتين . ويوضح الجدول ٧ القيم  $t$  والتي تحدد فوقها ( أى على يمينها ) مساحة معينة  $\alpha$  ، وذلك من أجل بعض القيم  $\alpha$  .

$$\alpha = 0.1, 0.05, 0.025, 0.01, 0.005$$

والجدول ٧ مصمم بصورة تختلف عن جدول التوزيع الطبيعي IV . فالسطر الأول يوضح المساحات المختلفة  $\alpha$  أما العمود الأيسر فيوضح درجات الحرية . ومن المألوف أن نرمز بـ  $T_{\alpha}$  للقيمة  $t$  التي تحدد فوقها ( أى على يمينها ) مساحة قدرها  $\alpha$  . لذلك فإن قيمة  $t$  بـ 15 درجة من الحرية والتي تحدد فوقها مساحة قدرها 0.005 هي القيمة  $t = 2.947$  . وبما أن التوزيع  $t$  متناظر حول الوسط المساوى للصفر ، لذلك فإننا نجد أن  $t_{1-\alpha} = -t_{\alpha}$  . وهذه هي قيمة  $t$  التي تحدد لنا مساحة على يمينها قدرها  $1 - \alpha$  ، ولذلك فإن قيمة  $t$  التي تحدد على يسارها مساحة قدرها  $\alpha$  تساوى ناقص القيمة  $t$  التي تحدد على يمينها نفس المساحة . لاحظ الشكل (٥،١٢) .



الشكل (٥،١٢)

نلاحظ في التوزيع  $t$  أنه من أجل 15 أن :

$$t_{0.995} = -t_{0.005} = -2.947$$

وهذا يعنى أن القيمة  $t$  من أجل عينة عشوائية حجمها  $n = 16$  ومختارة من مجتمع طبيعي ستقع بين القيمتين 2.947، -2.947 - باحتمال قدره 99% .

وبالضبط فإن 99% من التوزيع  $t$  بـ  $(n - 1)$  درجة من الحرية سيقع بين النقطتين  $t_{0.005}$  -  $t_{0.005}$  ، لذلك فإن القيمة  $t$  الواقعة قبل  $t_{0.005}$  - أو أعلى النقطة  $t_{0.005}$  ستجعلنا نميل إلى الاعتقاد في أن حادثاً نادراً سيقع ، أو أن افتراضاً حول قيمة  $\mu$  قد ارتكبنا فيه خطأ .

إن اهتمامنا بـ  $\mu$  سيحدد لنا طول المجال الذى نقبل فيه قيم  $t$  . وبعبارة أخرى إذا كنا نجهل قيمة الوسط الحقيقى والمختلفة بشكل تافه عما ندعيه أن تكون ، فعلينا اختيار مجال بعرض من  $t_{0.01}$  - إلى  $t_{0.01}$  ، وعلى  $t$  أن تقع عندئذ في هذا المجال . هذا وأن قيمة  $t$  الواقعة بالقرب من أحد أطراف المجال ستقودنا إلى الاعتقاد بأن فرضيتنا حول القيمة  $\mu$  صحيحة . كذلك من المرجح أن تكون بعض القيم القريبة من  $\mu$  ممثلة للقيمة الحقيقية ، وإذا كانت  $\mu$  معلومة بدرجة عالية من الدقة ، فيجب علينا عندئذ أن نستخدم أى مجال قصير مثل  $(t_{0.05}, t_{0.05})$  ، وفي هذه الحالة فإن قيمة  $t$  الواقعة خارج المجال سترسنا نعتقد أن القيمة المفروضة لـ  $\mu$  فيها خطأ ، في حين أن هذا صحيح بكل معنى الكلمة .

### مثال (٥, ١٩)

تدعى شركة معينة لإنتاج اللبمبات الكهربائية أن لبامها ستحترق عند مضى 700 ساعة في المتوسط على عملها . وللتأكد من هذا الإدعاء يقوم مهندسو هذه الشركة بإشعال 30 لمبة كل شهر ويقتنعوا بصحة إدعائهم إذا كانت القيمة المحسوبة لـ  $t$  واقعة في المجال  $(t_{0.05}, t_{0.05})$  . ما هو القرار النهائى الممكن استنتاجه من خلال عينة مسحوبة من إنتاج المصنع إذا كان لها وسط  $\bar{X} = 715$  ساعة ، وانحراف معيارى  $S = 38$  ساعة ؟ افرض أن توزيع احتراق لمبة توزيع قريب من التوزيع الطبيعي .

### الحل

من الجدول  $v$  نجد أن  $t_{0.05} = 1.699$  من أجل  $n - 1 = 29$  درجة من الحرية . لذلك فإن مهندسوا الشركة يقتنعون بصحة فرضيتهم إذا قدمت عينة عشوائية حجمها  $n = 30$  من إنتاج الشركة قيمة لـ  $t$  محصورة في المجال  $(-1.699, 1.699)$  . فإذا كان  $\mu = 700$  فعندئذ نجد أن :

$$t = \frac{715 - 700}{38 / \sqrt{30}} = 2.162$$

وهذه القيمة ستكون أكبر من 1.699 إن احتمال الحصول على قيمة  $\mu$  أكبر أو يساوى من 2.162 يساوى تقريبا 0.02 وإذا كانت  $\mu > 700$  فعندئذ ستكون قيمة  $\mu$  المحسوبة من خلال عينة معقولة بشكل أكبر . لذلك فإنه من المرجح أن تدعى الشركة أن إنتاجها من اللبمبات سيكون أفضل مما اعتقدت .

### (٥,٨) التوزيع F

إن أحد التوزيعات في التطبيقات الإحصائية هو التوزيع F . ويعرف التوزيع F على أنه يمثل نسبة متغيرين عشوائيين ( من نوع كاي - مربع ) مستقلان ، وكل متغير بدرجة حرية معينة . لذلك فإننا نكتب :

$$F = \frac{U / v_1}{V / v_2}$$

حيث يمثل  $U, V$  متغيرين عشوائيين مستقلين لهما توزيع كاي - مربع بعدد من درجات الحرية مساو لـ  $v_1, v_2$  على الترتيب . سنبحث فيما يلي عن الكثافة الاحتمالية للمتغير F .

### نظرية (٥,٩٨)

بفرض أن  $U$  و  $V$  متغيران عشوائيان مستقلان لهما توزيع كاي - مربع بعدد من درجات الحرية  $v_1, v_2$  على الترتيب . عندئذ يكون التوزيع الاحتمالي للمتغير :

$$F = \frac{U / v_1}{V / v_2}$$

من الشكل :

$$h(f) = \begin{cases} \frac{[\Gamma(v_1 + v_2)/2] (v_1/v_2)^{v_1/2}}{\Gamma(\frac{v_1}{2}) \cdot \Gamma(\frac{v_2}{2})} \cdot \frac{f^{v_1/2-1}}{(1 + \frac{v_1 f}{v_2})^{\frac{v_1+v_2}{2}}} & 0 < f < \infty \\ 0 & \text{فيما عدا ذلك} \end{cases}$$

تدعى الدالة  $h(f)$  بدالة كثافة التوزيع F بدرجتى الحرية  $v_1 - v_2$



## البرهان

من المعلوم أن التوزيع الاحتمالي المشترك للمتغيرين المستقلين  $u$  و  $v$  يعطى بالعلاقة :

$$(u, v) = f_1(u) \cdot f_2(v)$$

حيث يمثل كلا من  $f_1(u), f_2(v)$  كثافة المتغيرين  $u, v$  على الترتيب ، لذلك فإن :

$$\begin{aligned} \phi(u, v) &= \frac{1}{2^{\frac{n}{2}} \Gamma(\frac{n}{2})} \cdot u^{\frac{v_1}{2}-1} \cdot e^{-\frac{u}{2}} \frac{1}{2^{\frac{n}{2}} \Gamma(\frac{n}{2})} v^{\frac{v_2}{2}-1} \cdot e^{-\frac{v}{2}} \\ &= \begin{cases} \frac{1}{2^{\frac{v_1+v_2}{2}} \Gamma(\frac{v_1}{2}) \cdot \Gamma(\frac{v_2}{2})} u^{\frac{v_1}{2}-1} \cdot v^{\frac{v_2}{2}-1} e^{-\frac{1}{2}(u+v)} & 0 < u < \infty \\ & 0 < v < \infty \end{cases} \\ & \quad 0 \quad \text{فيما عدا ذلك :} \end{aligned}$$

لنعرف الآن متغيرا جديدا  $w = v$  . نلاحظ أن الحل العكسي للعلاقين  $f = \frac{U/v_1}{V/v_2}$  ،  $w = v$  هما  $w = V$  ،  $u = \frac{v_1}{v_2} f \cdot w$  ومنها نحصل على معين جاكوبى :

$$J = \begin{vmatrix} \frac{v_1}{v_2} w & \frac{v_1}{v_2} f \\ 0 & 1 \end{vmatrix} = \frac{v_1}{v_2} w$$

نلاحظ أن التطبيق بين مجموعة النقاط  $\{(u, v) | 0 < u < \infty, 0 < v < \infty\}$  ومجموعة النقاط  $\{(f, w) | 0 < f < w, 0 < w < \infty\}$  هو تطبيق واحد لواحد . وباستخدام النظرية (٥، ٤) نجد أن التوزيع الاحتمالي المشترك للمتغيرين  $W, F$  هو من الشكل :

$$g(f, w) = \begin{cases} \frac{w^{v_2/2-1}}{2^{(v_1+v_2)/2} \Gamma(\frac{v_1}{2}) \Gamma(\frac{v_2}{2})} \left( \frac{v_1}{v_2} f w \right)^{v_1/2-1} e^{-\frac{w}{2} \left[ \frac{v_1}{v_2} f + 1 \right]} & 0 < f < \infty \\ & 0 < w < \infty \end{cases}$$

فيما عدا ذلك :

أما توزيع المتغير F فنحسبه بواسطة التوزيع الهامشي :

$$h(f) = \int_0^{\infty} g(f, w) dw$$

$$= \frac{1}{2^{\frac{v_1+v_2}{2}} \Gamma(\frac{v_1}{2}) \cdot \Gamma(\frac{v_2}{2})} \int_0^{\infty} w^{\left(\frac{v_1+v_2}{2} - 1\right)} \cdot e^{-\frac{w}{2}\left[\frac{v_1}{v_2}f + 1\right]} dw$$

وباجراء تبديل في المتحول بالعلاقة :  $Z = \frac{w}{2} \left[ \frac{v_1}{v_2} f + 1 \right]$  نجد أن :

$$dw = \frac{2}{\left(\frac{v_1}{v_2} f + 1\right)} dz$$

ومنه :

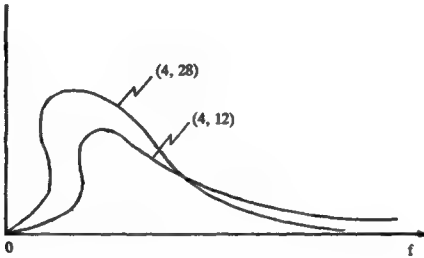
$$h(f) = \frac{\left(\frac{v_1}{v_2}\right)^{\frac{v_1}{2} - 1} f^{\frac{v_1}{2} - 1}}{2^{\frac{v_1+v_2}{2}} \cdot \Gamma(\frac{v_1}{2}) \cdot \Gamma(\frac{v_2}{2})}$$

$$\int_0^{\infty} \left( \frac{2Z}{\frac{v_1 f}{v_2 + 1}} \right)^{\frac{v_1+v_2}{2} - 1} \cdot e^{-Z} \frac{2}{\left(\frac{v_1 f}{v_2 + 1} + 1\right)} dz$$

$$= \frac{\left(\frac{v_1}{v_2}\right)^{\frac{v_1}{2} - 1} f^{\frac{v_1}{2} - 1}}{\Gamma(\frac{v_1}{2}) \cdot \Gamma(\frac{v_2}{2}) \cdot \left(1 + \frac{v_1 f}{v_2}\right)^{\frac{v_1+v_2}{2}}} \int_0^{\infty} z^{\frac{v_1+v_2}{2} - 1} e^{-z} dz$$

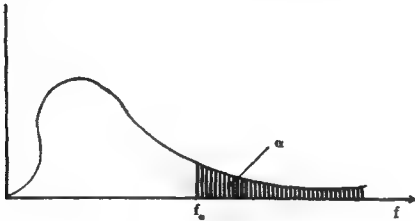
$$= \begin{cases} \frac{\Gamma(\frac{v_1+v_2}{2}) \left(\frac{v_1}{v_2}\right)^{\frac{v_1}{2} - 1} f^{\frac{v_1}{2} - 1}}{\Gamma(\frac{v_1}{2}) \cdot \Gamma(\frac{v_2}{2}) \left(1 + \frac{v_1 f}{v_2}\right)^{\frac{v_1+v_2}{2}}} : 0 < f < \infty \\ 0 : \text{فيما عدا ذلك} \end{cases}$$

نلاحظ في عبارة  $F$  أن درجة حرية المتغير الأول  $u$  قد وردت في بسط الكسر . كما أن درجة حرية المتغير الثاني  $v$  قد وردت في مقام الكسر . وهذا يعنى أن منحني توزيع المتغير  $F$  لا يتعلق بدرجة حرية المتغيرين  $u$  و  $v$  وإنما يتعلق أيضا بالترتيب الذي تبدأ فيه . ويوضح الشكل (٥, ١٣) منحنيين للتوزيع  $F$  من أجل درجات الحرية (4, 12) و (4, 28) .



الشكل (٥, ١٣)

لنفرض أن  $f_\alpha$  تمثل قيمة خاصة للمتغير العشوائي  $F$  ، والتي تحدد لنا على يمينها مساحة قدرها  $\alpha$  يوضحها القسم المظلل على الشكل (٥, ١٤) الموضح أدناه .



الشكل (٥, ١٤)

ويوضح لنا الجدول VII القيم  $f_{\alpha}$  من أجل  $\alpha = 0.05$  ،  $\alpha = 0.01$  وذلك من أجل عدد من الأزواج ( ١ ، ٢ ) . نلاحظ أن قيمة  $f$  بـ (4, 12) درجات من الحرية والتي تحدد لنا على يمينها مساحة قدرها 0.05 هي القيمة  $f_{0.05} = 3.26$  . وتوضح لنا النظرية التالية طريقة لاستخدام الجدول VII في حساب قيم  $f_{0.95}$  ،  $f_{0.99}$  .

### نظرية (٥, ١٩)

بفرض أن  $f_{\alpha}$  ( ١ ، ٢ ) تمثل القيمة  $f_{\infty}$  بدرجتي الحرية ( ١ ، ٢ ) عندئذ :

$$f_{1 - \alpha}(p_1, p_2) = \frac{1}{f_{\alpha}(p_1, p_2)}$$

وهكذا فإن قيمة  $f$  بدرجتي الحرية (4, 12) والتي تحدد لنا على يمينها مساحة قدرها 0.95 هي :

$$f_{0.95}(4, 12) = \frac{1}{f_{0.05}(12, 4)} = \frac{1}{5.91} = 0.169$$

نفرض أننا سحبنا عيتين عشوائيتين حجمهما على الترتيب  $n_1$  ،  $n_2$  من مجتمعين طبيعيين بالتباينين  $\sigma_1^2$  ،  $\sigma_2^2$  على التالى . من النظرية (٥, ١٦) نعلم أن لكل من المتغيرين :

$$\chi_1^2 = \frac{(n_1 - 1) S_1^2}{\sigma_1^2}$$

$$\chi_2^2 = \frac{(n_2 - 1) S_2^2}{\sigma_2^2}$$

توزيعا من نوع كاي - مربع بـ  $\nu_1 = n_1 - 1$  ،  $\nu_2 = n_2 - 1$  على الترتيب درجات من الحرية . علاوة على ذلك ، بما أن العينات عشوائية ، لذلك فإنها ستكون مستقلة ، وباستخدام النظرية (٥, ١٨) بعد فرض أن :

$$\chi_1^2 = U, \chi_2^2 = V$$

نجد النظرية التالية :

### نظرية (٥,٢٠)

بفرض أن  $S_1^2, S_2^2$  يمثلان تباينى عييتين مستقلتين حجماهما على الترتيب  $n_1, n_2$  مأخوذتين من مجتمعين طبيعيين بالتباينين  $\sigma_1^2, \sigma_2^2$  فعندئذ يكون للمتغير :

$$F = \frac{S_1^2 / \sigma_1^2}{S_2^2 / \sigma_2^2} = \frac{\sigma_2^2 S_1^2}{\sigma_1^2 S_2^2}$$

توزيعا من نوع F بعدد من درجات الحرية مساوى لـ :

$$(v_1, v_2) = (n_1 - 1, n_2 - 1)$$

في الفصل السادس سنستخدم النظرية (٥,٢٠) لإيجاد ثقة لتباينى مجتمعين طبيعيين ،

يستخدم التوزيع F على نطاق واسع في طرق الإحصاء . ونلاحظ أن  $t^2 = \frac{Z^2}{V/v}$  هو نسبة  $\chi^2(1)$  إلى  $\chi^2(v)$  أى أن توزيع المتغير  $[t(v)]^2$  هو بالتعريف  $F(1, v)$  ، وهذا يعنى أن التوزيع  $F(1, v)$  يكافئ مربع التوزيع  $t(v)$  .

## تمارين محلولة

## تمرين (١)

بفرض  $X$  متغيراً عشوائياً حدانياً . توزيعه الاحتمالي من الشكل :

$$f(x) = \begin{cases} \left(\frac{3}{x}\right) \cdot \left(\frac{2}{5}\right)^x \left(\frac{2}{5}\right)^{3-x} & : x = 0, 1, 2, 3 \\ 0 & : \text{فيما عدا ذلك} \end{cases}$$

ما هو التوزيع الاحتمالي للمتغير الجديد  $Y = X^2$  ؟

## الحل

نلاحظ أن المتغير المفروض  $X$  منقطع وقيم جميعها موجبة ، لذلك فالتقابل  $Y = X^2$  هو تقابل واحد لواحد بين مجموعة قيم  $x$  وقيم  $y$  ، وحسب النظرية (٥،١) نستطيع أن نكتب بعد حل المعادلة  $y = x^2$  بالنسبة لـ  $x$  .

$$g(y) = f\{\sqrt{y}\} = \begin{cases} \left(\frac{3}{\sqrt{y}}\right) \left(\frac{2}{5}\right)^{\sqrt{y}} \left(\frac{2}{5}\right)^{3-\sqrt{y}} & : y = 0, 1, 4, 9 \\ 0 & : \text{فيما عدا ذلك} \end{cases}$$

## تمرين (٢)

بفرض أن التوزيع الاحتمالي المشترك  $X_1, X_2$  هو من الشكل :

$$f(x_1, x_2) = \left(\frac{2}{x_1, x_2, 2-x_1-x_2}\right) \cdot \left(\frac{1}{4}\right)^{x_1} \cdot \left(\frac{1}{3}\right)^{x_2} \cdot \left(\frac{5}{12}\right)^{2-x_1-x_2}$$

وذلك من أجل  $x_1 + x_2 \leq 2$  ,  $x_2 = 0, 1, 2$  ,  $x_1 = 0, 1, 2$  وصفر فيما عدا ذلك :

ما هو التوزيع المشترك للمتغيرين  $Y_1 = X_1 + X_2$  و  $Y_2 = X_1 - X_2$  ؟

**الحل**

من الواضح أن التطبيقين  $J_1$  و  $J_2$  هما تطبيقان واحد لواحد بين مجموعة النقاط  $(y_1, y_2)$  ومجموعة النقاط  $(Y_1, Y_2)$  وحسب النظرية (٥, ٢٠) نكتب [بعد حل المعادلتين  $y_2 = x_1 - x_2, y_1 = x_1 + x_2$  حلا عكسيا بالنسبة لكل من  $x_1$  و  $x_2$  نجد أن :

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \frac{y_1 + y_2}{2} \\ \frac{y_1 - y_2}{2} \end{bmatrix}$$

$$g(y_1, y_2) = f \left[ \frac{y_1 + y_2}{2}, \frac{y_1 - y_2}{2} \right]$$

$$g(y_1, y_2) = \begin{cases} \left( \frac{2}{y_1 + y_2, \frac{y_1 - y_2}{2}, 2 - y_1} \right) \left( \frac{1}{4} \right)^{\frac{y_1 + y_2}{2}} \left( \frac{1}{3} \right)^{\frac{y_1 - y_2}{2}} \left( \frac{1}{11} \right)^{2 - y_1} & \begin{matrix} y_1 = 0, 1, 2 \\ y_2 = -1, 0, 1 \end{matrix} \\ 0 & \text{فيما عدا ذلك :} \end{cases}$$

وهكذا نجد أن :

$$g(y_1, y_2) = \begin{cases} \left( \frac{2}{y_1 + y_2, \frac{y_1 - y_2}{2}, 2 - y_1} \right) \left( \frac{121}{300} \right)^{\frac{y_1}{2}} \cdot \left( \frac{3}{4} \right)^{\frac{y_2}{2}} \left( \frac{5}{11} \right)^2 : & \begin{matrix} y_1 = 0, 1, 2, y_2 = -1, 0, 1 \\ y_2 \leq y_1, y_1 + y_2 \text{ are even} \end{matrix} \\ 0 & \text{فيما عدا ذلك :} \end{cases}$$

**تمرين (٣)**

بفرض أن  $X_1, X_2$  يمثلان متغيرين عشوائيين منقطعين بالتوزيع الاحتمالي المشترك :

$$f(x_1, x_2) = \begin{cases} \frac{x_1 x_2}{18} & \begin{matrix} x_1 = 1, 2 \\ x_2 = 1, 2, 3 \end{matrix} \\ 0 & \text{فيما عدا ذلك :} \end{cases}$$

أوجد التوزيع الاحتمالي للمتغير  $Y = X_1 \cdot X_2$

**الحل**

لنفترض أولاً عن التوزيع الاحتمالي المشترك للمتغيرين  $Y, X_2$ . فحسب النظرية (٥, ٢) وبملاحظة أن كلا من التقابلين  $Y, X_2$  هو واحد لواحد فإننا نجد أن :

$$g(y, x_2) = f \left[ \frac{y}{x_2}, x_2 \right]$$

$$= \begin{cases} \frac{\frac{y}{x_2} \cdot x_2}{18} & : y = 1, 2, 3, 4, 6 \\ & x_2 = 1, 2, 3, y/x_2 = 1, 2 \\ 0 & : \text{فيما عدا ذلك} \end{cases}$$

وبالجمع على قيم  $x_2 = 1, 2, 3$  نجد التوزيع الهامشي للمتغير  $y$  ومنه :

$$g(y) = \begin{cases} \frac{y}{18} & : y = 1, 3, 4, 6 \\ \frac{2y}{18} & : y = 2 \\ 0 & : \text{فيما عدا ذلك} \end{cases}$$

**تمرين (٤)**

إذا علمت أن للمتغير العشوائي المستمر  $X$  توزيعاً طبيعياً بالوسط  $\mu = 0$  والتباين  $\sigma^2$  فما هي دالة الكثافة للمتغير  $Y = aX^2$  وحيث أن  $a > 0$  ؟

**الحل**

لقد برهنا في التمرين (٥, ٥) أنه إذا كان للمتغير  $x$  توزيعاً طبيعياً بالمتغيرين  $\mu, \sigma$  كان للمتغير  $\left( \frac{x - \mu}{\sigma} \right)^2$  توزيعاً من نوع كاي - مربع بدرجة واحدة من الحرية وبالعودة إلى التمرين المذكور وبفرض أن  $\mu = 0$  نجد أن للمتغير  $Z = \frac{X^2}{\sigma^2}$  توزيعاً يشبه توزيع كاي - مربع بدرجة واحدة من الحرية أي أن للمقدار  $X^2$  توزيعاً من الشكل :



$$f_{\chi^2}(z) = \begin{cases} \frac{1}{\sqrt{2\pi z} \cdot \sigma} e^{-\frac{1}{2} \frac{z}{\sigma^2}} & : z > 0 \\ 0 & : \text{فيما عدا ذلك} \end{cases}$$

ومن المعلوم حسب النظرية (٥,٣) أن :

$$f_Y(y) = f_{\chi^2}(y/a) \cdot \left[ \frac{d(x^2)}{dy} \right]$$

محسوبا عند علاقة التحويل  $x = \frac{y}{a}$

لذلك فإن :

$$f_Y(y) = \begin{cases} \frac{1}{\sqrt{2\pi a \cdot y} \sigma} e^{-\frac{1}{2} \frac{y}{a \cdot \sigma^2}} & : y > 0 \\ 0 & : \text{فيما عدا ذلك} \end{cases}$$

تمرين (٥)

بفرض أن دالة كثافة الثنائية  $(x, y)$  هي من الشكل :

$$f(x, y) = |J| \cdot e^{-\frac{1}{2} \left[ \frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} \right]}$$

$$= \frac{1}{\sigma_x \sigma_y 2\pi} e^{-\frac{1}{2} \left[ \frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} \right]}$$

فما هي دالة كثافة المتغير  $x$  ؟  $y$

الحل

لنبحث أولا عن دالة الكثافة للثنائية  $(x, y)$

فحسب النظرية (٥,٤) يمكن أن نكتب :

$$f(z, y) = f(x + y, y) \cdot |J|$$

$$= \frac{1}{\sigma_x \sigma_y 2\pi} e^{-\frac{1}{2} \left[ \frac{(x+y)^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} \right]}$$

ذلك لأن :

$$|J| = \begin{vmatrix} \frac{\partial x}{\partial z} & \frac{\partial x}{\partial y} \\ \frac{\partial y}{\partial z} & \frac{\partial y}{\partial y} \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ \frac{\partial y}{\partial z} & 1 \end{vmatrix} = 1$$

وبعد مكاملة الدالة الأخيرة بالنسبة لـ  $y$  نجد دالة الكثافة للمتغير  $Z$  هي .

$$f(z) = \frac{1}{\sigma_x \sigma_y 2\pi} \int_{y=-\infty}^{+\infty} \frac{1}{e^2} \left| \frac{(z+y)^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} \right| dy$$

$$= \frac{e^{-\frac{1}{2} \frac{z^2}{\sigma_x^2}}}{\sigma_x \sigma_y 2\pi} \int_{y=-\infty}^{+\infty} \frac{1}{e} \left| y^2 \left( \frac{\sigma_x^2 + \sigma_y^2}{\sigma_x^2 \cdot \sigma_y^2} \right) + 2yz \sigma_y^2 \right| / \sigma_x^2 \cdot \sigma_y^2$$

لنتم لمربع كامل فنجد أن :

$$= \frac{e^{-\frac{z^2}{2} \left( \frac{1}{\sigma_x^2} + \frac{1}{\sigma_y^2} \right)}}{\sigma_x \sigma_y \sqrt{2\pi}}$$

$$\sigma_z = \sqrt{\sigma_x^2 + \sigma_y^2} \quad \text{حيث إن :}$$

تموين (٦)

يفرض أن التوزيع المشترك للمتغيرين العشوائيين المستمرين  $X, Y$  هو من الشكل :

$$f(x, y) = \frac{1}{2\pi \cdot \sigma^2} e^{-\frac{x^2 + y^2}{2\sigma^2}}$$

وبفرض أن  $Y = R \sin \theta$ ,  $X = R \cos \theta$

فما هو التوزيع المشترك للمتغيرين  $R, \theta$  ؟

## الحل

حسب النظرية (٥,٤) نكتب أن دالة الكثافة للمتغيرين  $R$ , هي من الشكل :

$$f(r, \theta) = f(x, y) \cdot \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial y}{\partial r} \\ \frac{\partial x}{\partial \theta} & \frac{\partial y}{\partial \theta} \end{vmatrix}$$

محسوبة عند  $y = r \sin \theta$ ,  $x = r \cos \theta$

ومنه :

$$f(r, \theta) = \frac{r}{2 \cdot \pi \cdot \sigma^2} e^{-\frac{1}{2\sigma^2} r^2} : \begin{matrix} 0 < r < \infty \\ 0 \leq \theta \leq 2\pi \end{matrix}$$

## تمرين (٧)

برهن أن الدالة المولدة للعزوم لمتغير عشوائي هندسي هي من الشكل :

$$M_X(t) = \frac{p e^t}{1 - q e^t}$$

## البرهان :

من المعلوم أن التوزيع الاحتمال للمتغير الهندسي يعطى بالجدول :

X	1	2	3	...	x	...
f(x)					$p \cdot q^{x-1}$	

وحسب التعريف (٥,١) نجد أن :

$$\begin{aligned} M_X(t) &= E \{ e^{tx} \} = \sum_{x=1}^{\infty} e^{tx} \cdot f(x) \\ &= \sum_{x=1}^{\infty} e^{tx} \cdot p \cdot q^{x-1} \\ &= \frac{p}{q} \sum_{x=1}^{\infty} (e^t \cdot q)^x \end{aligned}$$

$$= \frac{P}{q} \cdot \left[ \sum_{x=0}^{\infty} (e \cdot q)^x \right] \cdot q \cdot \frac{1}{e}$$

$$\sum_{x=0}^{\infty} (qe^t)^x = \frac{1}{1 - qe^t} \quad \text{ومن الواضح أن } |e^t \cdot q| \leq 1$$

لذلك فإن :

$$M_x(t) = \frac{p e^t}{1 - qe^t}$$

تمارين (٨)

أوجد الدالة المولدة للعزوم لتغير عشوائي بواسوني  $X$  بالوسط  $\mu$ .

الحل

نعلم أن التوزيع الاحتمالي لتغير بواسون بالوسط  $\mu$  هو من الشكل :

X	0	1	2	...	x	...
f(x)					$e^{-\mu} \frac{\mu^x}{x!}$	

وبالعودة إلى التعريف (٥, ١) فإننا نجد أن :

$$\begin{aligned} M_x(t) &= \sum_{x=0}^{\infty} e^{tx} \cdot e^{-\mu} \cdot \frac{\mu^x}{x!} \\ &= e^{-\mu} \cdot \sum_{x=0}^{\infty} \frac{(\mu \cdot e^t)^x}{x!} \end{aligned}$$

ومن المعلوم أن :

$$\sum_{x=0}^{\infty} \frac{\alpha^x}{x!} = e^{\alpha}$$

إذن:

$$\begin{aligned} M_X(t) &= e^{-\mu} \cdot e^{\mu e^t} \\ &= e^{(e^t - 1)\mu} \end{aligned}$$

وهي الدالة المولدة للعزوم لتغير بواسوني بالوسط  $\mu$ .

تمرين (٩)

مستخدما الدالة المولدة للعزوم في التمرين (٥,٨) بين أن توقع وتباين متغير كاي - مربع بـ  $p$  درجة حرية هما على الترتيب .

البرهان

من التمرين (٥,٨) نجد أن :

$$M_X(t) = \frac{1}{(1 - 2t)^{p/2}}$$

وحسب النظرية (٥,٦) نعلم أن :

$$\mu = \left. \frac{dM_X(t)}{dt} \right|_{t=0}, \quad \mu_2 = \left. \frac{d^2 M_X(t)}{dt^2} \right|_{t=0}$$

وباشتقاق الدالة السابقة مرتين بالنسبة لـ  $t$  فإننا نجد أن :

$$\frac{dM_X(t)}{dt} = -\frac{p}{2} (-2) (1 - 2t)^{-\frac{p}{2}-1} = \frac{p}{(1 - 2t)^{\frac{p}{2} + 1}}$$

$$\frac{d^2 M_X(t)}{dt^2} = -\left(\frac{p}{2} + 1\right) (-2) (1 - 2t)^{-\frac{p}{2}-2}$$

ويوضع  $t = 0$  في العلاقتين السابقتين فإننا نجد :

$$\mu = \nu, \quad \mu_2 = 2\nu \left( \frac{\nu}{2} + 1 \right)$$

$$\sigma^2 = \mu_2 - \mu^2 = 2\nu \left( \frac{\nu}{2} + 1 \right) - \nu^2 = 2\nu \quad \text{ومنه :}$$

تمرين (١٠)

استخدم الجدول VI في إيجاد :

$$\nu = 18 \quad \frac{2}{0.01} \quad (١)$$

$$\nu = 29 \quad \frac{2}{0.975} \quad (٢)$$

$$\nu = 4 \quad \frac{2}{\alpha} \quad (٣) \quad \text{بحيث تتحقق المعادلة } P[X^2 < \chi^2_{\alpha}] = 0.99 \text{ وذلك من أجل } \nu = 4 :$$

الحل

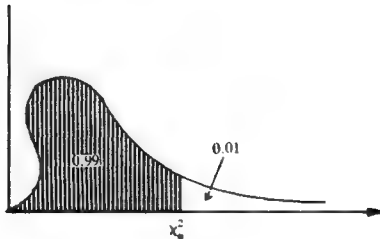
من الجدول VI نجد أن :

$$\chi^2_{0.01} = 34.805 \quad : \quad \nu = 18$$

$$\chi^2_{0.975} = 16.047 \quad : \quad \nu = 29$$

ولإيجاد قيمة  $\frac{2}{\alpha}$  التي نحدد على يسارها مساحة قدرها 0.99 علينا أن نبحث في الجدول المذكور عن  $\frac{2}{\alpha}$  التي نحدد على يمينها مساحة قدرها 0.01 من أجل  $\nu = 4$ . لذلك فإننا نجد أن قيمة  $\frac{2}{\alpha}$  المطلوبة والمحققة للعلاقة :

$$P[X < \chi^2_{\alpha}] = 0.99$$

هي  $\chi^2_{0.01}$  ومنه  $\chi^2_{\alpha} = 13.277$  انظر الشكل المرفق :

## تقريب (١١)

احسب احتمال أن يكون تباين عينة  $S^2$  حجمها  $n = 25$  مسحوبة من مجتمع طبيعي بالتباين  $\sigma^2 = 6$

(١) أكبر من 9.1

(٢) بين القيمتين (3.462, 10.745)

## الحل

أولاً — نعلم أن :

$$P [S^2 > 9.1] = P \left[ \frac{(n-1) S^2}{\sigma^2} > \frac{24}{6} (9.1) \right]$$

$$= P [\chi^2_{(24)} > 36.4]$$

لنبحث الآن عن المساحة التي يعدها العدد 36.4 على يمينه في التوزيع كاي - مربع بـ  $n = 24$  درجة من الحرية . من الجدول VI نجد أن هذا الاحتمال يمثل المساحة 0.05 .

ثانياً — نعلم أن :

$$P [3.462 < S^2 < 10.745] = P \left[ 13.848 < \frac{(n-1) S^2}{\sigma^2} < 42.98 \right]$$

$$= P [13.848 < \chi^2_{(24)} < 42.98]$$

$$= \chi^2_{13.848} - \chi^2_{42.98}$$

$$= 0.95 - 0.01$$

$$= 0.94$$

## تمارين (١٢)

باستخدام الجدول  $V$  أوجد :

$$(١) \quad t_{0.025} \text{ من أجل } \nu = 17$$

$$(٢) \quad t_{0.09} \text{ من أجل } \nu = 10$$

$$(٣) \quad t_{\alpha} \text{ بحيث تتحقق العلاقة } P[-t_{\alpha} < T < t_{\alpha}] = 0.90$$

وذلك من أجل  $\nu = 23$

## الحل

من الجدول  $V$  نجد أن :

$$t_{0.025} = 2.11 \quad , \quad \nu = 17$$

$$\text{ومن المعلوم أن } t_{\alpha} = -t_{1-\alpha} \text{ لذلك فإنه من أجل } 10$$

$$t_{0.99} = -t_{0.01} = -2.764$$

ثالثا — نلاحظ أن  $t_{\alpha}$  المطلوب حسابها يجب أن تحدد على يمينها مساحة قدرها 0.05 ، وذلك لأن توزيع  $t$  متناظر حول الوسط الذي يساوى صفرا . وهكذا نجد أن :

$$t_{0.05} = 1.714$$

وذلك من أجل  $\nu = 23$  .



## تمارين عامة

(١) ليكن  $X$  متغيراً عشوائياً منقطعاً بالتوزيع الاحتمالي :

$$f(x) = \begin{cases} \frac{1}{3} & : x = 1, 2, 3 \\ 0 & : \text{فيما عدا ذلك} \end{cases}$$

أوجد التوزيع الاحتمالي للمتغير  $Y = 3X - 3$

(٢) بفرض أن  $X$  يمثل متغيراً عشوائياً حدانياً بالتوزيع الاحتمالي :

$$f(x) = \begin{cases} \left(\frac{3}{x}\right) \left(\frac{2}{5}\right)^x \left(\frac{3}{5}\right)^{3-x} & : x = 0, 1, 2, 3 \\ 0 & : \text{فيما عدا ذلك} \end{cases}$$

أوجد التوزيع الاحتمالي للمتغير  $Y = X^3$

(٣) ليكن  $X_1, X_2$  متغيرين عشوائيين منقطعين بالتوزيع المشترك :

$$f(x_1, x_2) = \begin{cases} \left(\frac{1}{4}\right)^{x_1} \left(\frac{1}{3}\right)^{x_2} \left(\frac{5}{12}\right)^{2-x_1-x_2} & : x_1 = 0, 1, 2 \\ & : x_2 = 0, 1, 2 \\ 0 & : \text{فيما عدا ذلك} \end{cases}$$

ما هو التوزيع الاحتمالي المشترك للمتغيرين :

$$Y_1 = X_1 + X_2, \quad Y_2 = X_1^2 + X_1 \cdot X_2$$

(٤) بفرض أن للمتغير العشوائي المستمر  $X$  كثافة احتمالية من الشكل :

$$f(x) = \begin{cases} 1 & : 0 < x < 1 \\ 0 & : \text{فيما عدا ذلك} \end{cases}$$

برهن أن للمتغير  $Y = -2 \ln X$  توزيعاً من نوع كاي - مربع بدرجة حرة .

(٥) يتغير التيار  $I$  أمبير المتدفق في المقاومة  $R$  أوم وفقاً للتوزيع الاحتمالي :

$$f(i) = \begin{cases} 6i(1-i) & : 0 < i < 1 \\ 0 & : \text{فيما عدا ذلك} \end{cases}$$

فإذا علمت أن المقاومة تتغير بصورة مستقلة عن التيار وفقاً للتوزيع الاحتمالي :

$$g(r) = \begin{cases} 2r & : 0 < r < 1 \\ 0 & : \text{فيما عدا ذلك} \end{cases}$$

فما هو التوزيع الاحتمالي للقوة  $W = I^2.R$  واط ؟

(٦) بفرض أن التوزيع المنتظم للمتغير المنقطع  $X$  هو من الشكل :

$$f(x) = \begin{cases} \frac{1}{n} & : x = 1, 2, 3, \dots, n \\ 0 & : \text{فيما عدا ذلك} \end{cases}$$

أ — برهن أن الدالة المولدة للعزوم للمتغير  $X$  هي من الشكل :

$$M_X(t) = \frac{e^t(1 - e^{-t})}{n(1 - e^{-t})}$$

ب — باستخدام الدالة  $M_X(t)$  السابقة . احسب توقع وتباين المتغير  $X$  السابق .

(٧) استخدم الدالة المولدة للعزوم في التمرين (٦) في حساب توقع وتباين المتغير الهندسي .

(٨) إذا علمت أن الدالة المولدة للعزوم لمتغير بواسوني معين هي من الشكل :

$$M_X(t) = e^{4e^t - 1}$$

فأوجد :

$$P(\mu - 2\sigma < X < \mu + 2\sigma)$$

(٩) سجلنا عدد أيام الغياب لثمانية عشر طالباً في كلية الهندسة خلال العام الفائت فوجدنا أنها تشكل عينة عشوائية من الشكل :

$$1, 3, 4, 0, 4, 2, 3, 1, 2, 3, 0, 4, 1, 1, 1, 5, 1, 0$$

احسب وسط العينة .

(١٠) احسب تباين العينة 3,5,8,7,5,7

(١١) أوجد تباين العينة 6,10,16,14,10,14 بدون حساب .

(١٢) سحبنا عينة حجمها  $N = 16$  من مجتمع طبيعي وسطه  $\mu = 50$  وتباينه  $\sigma^2 = 25$  .  
أوجد احتمال أن يقع وسط العينة  $\bar{X}$  ضمن المجال :  
(  $\sigma_{\bar{X}}^2 - 1.9\mu_{\bar{X}}, \mu_{\bar{X}} + 0.4\sigma_{\bar{X}}$  )

وذلك بفرض أنه يمكن قياس وسط العينة بأى درجة من الدقة .

(١٣) بفرض أن لأطوال ألف طالب توزيعاً قريباً من التوزيع الطبيعي بوسط قدره 174.5 سنتيمتر وانحراف معياري قدره 6.9 سنتيمتر . لنسحب مئتي عينة مؤلفة من خمسة وعشرين طالباً وبشكل عشوائي من هذا المجتمع ولنسجل وسطاء العينات بالنسبة لأقرب عشر من السنتيمتر ما هو :

أ — توقع الوسط والانحراف المعياري لوسط توزيع المعاينة .

ب — عين عدد وسطاء العينات الواقعة في المجال [172.5, 175.8] .

ج — عين عدد وسطاء العينات الواقعة أسفل العدد 172 سنتيمتر .

(١٤) تجربة حدانية مؤلفة من  $n = 25$  اختباراً فيها  $p = 0.4$  . احسب  $P(8 \leq X \leq 11)$  وذلك :

أ — باستخدام جدول التوزيع الحداني ( الجدول II ) .

ب — باستخدام طريقة التقريب الطبيعي للتوزيع الحداني .

(١٥) وجد بالتجربة أن الفترة اللازمة لإنجاز اختبار للذكاء مخصص لطلبة كلية الهندسة يتوزع وفقاً للتوزيع الطبيعي بوسط  $\mu = 65$  دقيقة وانحراف معياري قدره  $\sigma = 10$  دقائق . فكم يجب أن يكون زمن الاختبار إذا أردنا إتاحة وقت كاف لـ 80% من الطلاب لإتمام الاختبار ؟

(١٦) يتوزع عمر نوع من الثلاجات الكهربائية وفقاً للتوزيع الطبيعي بوسط  $\mu = 3.1$  سنة وانحراف معياري  $\sigma = 1.2$  سنة . فإذا كانت الثلاجات مكفولة لمدة عام ، فما هي نسبة الثلاجات المباعة والتي ستضطر الشركة المنتجة لاستبدالها بثلاجات جديدة ؟

(١٧) بنشر الدالة الأسية  $e^{tx}$  حسب سلسلة ماكلوران ثم مكاملة كل حد من حدودها ، برهن أن :

$$M_x(t) = \int_{-\infty}^{+\infty} e^{tx} \cdot f(x) dx$$

$$= 1 + \mu_1 t + \mu_2 \frac{t^2}{2!} + \dots + \mu_v \frac{t^v}{v!} + \dots$$

(١٨) عيتان عشوائيتان حجمهما على الترتيب  $n_1 = 25$  ،  $n_2 = 36$  مأخوذتين من مجتمعين مختلفين بالوسطين  $\mu_1 = 80$  ،  $\mu_2 = 75$  وانحرافين معيارين  $\sigma_1 = 5$  ،  $\sigma_2 = 3$  على التوالي . احسب احتمال أن يكون وسط العينة  $\bar{X}_1$  أكبر من وسط العينة  $\bar{X}_2$  بـ 3.4 على الأقل ، وأقل من 5.9 .

(١٩) هل من الممكن الحصول على عينة حجمها  $n = 9$  ، وسطها  $\bar{X} = 24$  ، وانحرافها المعياري  $\sigma_{\bar{X}} = 4.1$  من مجتمع طبيعي تباينه مجهول ووسطه  $\mu = 20$  ؟

(٢٠) باستخدام الجدول VII احسب :

$$f_{0.05} \text{ من أجل } p_1 = 7 \quad p_2 = 19$$

$$f_{0.01} \text{ من أجل } p_1 = 24 \quad p_2 = 19$$

$$f_{0.95} \text{ من أجل } p_1 = 19 \quad p_2 = 24$$

$$f_{0.99} \text{ من أجل } p_1 = 28 \quad p_2 = 12$$

(٢١) أعلنت مديرية الإذاعة والتلفزيون أن 20% من المشاهدين يتابعون برنامجا معيناً . وقد وجد من خلال عينة عشوائية مؤلفة من ألف مشاهد ، أن 150 شخصا من بينهم يتابعون البرنامج المذكور فهل تقدم هذه المعلومات دلالة كافية لتأييد ما أعلنته مديرية الإذاعة والتلفزيون ؟

(٢٢) تدعى شركة لإنتاج السجائر أن متوسط القطران في سجائرها 18.3 مليغرام . اخترنا عينة عشوائية من إنتاج هذه الشركة حجمها  $n = 8$  سيجارة ، وحللناها وسجلنا وسط القطران الموجود فيها فوجدنا أنه :

20, 17, 21, 19, 22, 21, 20, 16

فهل توافق إدعاء الشركة السابق أم لا ؟

(٢٣) سحبنا عيتين عشوائيتين مستقلتين حجمهما على الترتيب  $n_1 = 25$  ،  $n_2 = 31$  من مجتمعين طبيعيين بالتباينين  $\sigma_1^2 = 10$  ،  $\sigma_2^2 = 15$  على الترتيب . احسب :

$$P(S_1^2 / S_2^2 < 1.26)$$



# الفصل السادس

## نظرية التقدير

■ مقدمة ■ طرق التقدير الكلاسيكية ■ تقدير الوسط ■ تقدير  
فروق وسطين ■ تقدير  $P$  في المجتمع الحداثي ■ تقدير الفرق بين نسبي  
مجموعتين حداثيتين ■ تقدير التباين ■ تقدير نسبة تباينين ■ نماذج  
عائلة ■ نماذج عامة .





## (١, ٦) مقدمة

يواجه الفرد خلال حياته اليومية حالات تتطلب منه القيام بتنبؤات حول المستقبل . فنجد مثلاً أن شركة صناعية للمبات الكهربائية تستخدم نتائج تجربة معينة لاستقراء ما إذا كان نوعاً جديداً من اللببات أكثر جودة من نوع آخر . كما تهتم كل حكومة بالتنبؤ بعدد الطلاب في مختلف المراحل خلال السنوات العشر القادمة . ويرغب مهندسو شركة لإنتاج الأسلحة في معرفة ما إذا كان الفولاذ المستورد من بلد معين أكثر مقاومة لتغيرات الحرارة من نوع آخر . ومن حسن الحظ أنه يمكن أن نتخذ قراراً ونقوم بتنبؤ حول الشيء المدروس في كل حالة من خلال بعض المعلومات المتوفرة لنا ، والتي ندعوها بالملاحظات أو المعلومات الإحصائية . وفي العديد من الحالات تكون المعلومات المتوفرة لدى المدارس لحالة معينة غريزية ، منسجمة في ظاهرها ، ومن نواح عديدة على درجة كبيرة من الوضوح بحيث لا يكون التنبؤ المتخذ بضائة أفضل بكثير من مجرد التخمين الشخصي . ومن جهة أخرى نجد أن التحليل الشخصي للمعلومات من قبل المهندسين مثلاً يؤدي غالباً إلى آراء متعارضة حول النتائج المستخلصة لتجربة معينة . وفي الوقت الذي يتفق فيه العديد من الناس بالشعور بقدرتهم الذاتية على القيام باستقراءات جيدة ، إلا أن التجربة تدل على أنه ليس باستطاعة الإنسان أن يجهد فكره بكمية كبيرة من الأرقام وتحليل وموازنة القليل من المعلومات للوصول إلى استقراء جيدة . لهذا يصبح وضع قوانين وأدوات للاستقراء أمراً مرغوباً به وهذا هو هدف الإحصاء .

إن هدف الإحصاء هو القيام باستقراءات حول المجتمع المدروس بدءاً من معلومات تقدمها العينة . وإلى المدى الذي تتميز فيه المجتمعات المدروسة بمقاييس وصفية رقمية تدعى بالوسطاء ، لذلك فالاستقراء الاحصائي يهتم بتوفير استقراءات جيدة حول هؤلاء الوسطاء . ومن الأمثلة على هؤلاء الوسطاء المتوسط والنشت .

والقيام باستقراء حول وسيط يمكن أن ينجز بطريقتين . فيمكن أن نقوم بتقدير الوسيط أى التنبؤ بقيمته أو أن نتخذ قراراً يتعلق بقيمة الوسيط . ونلاحظ أنه لا بد من وجود مقياس بجودة كل طريقة بحيث يمكن مقارنة هذه الطرق ببعضها والقيام بمفاضلة فيما بينها . هذا بالإضافة إلى أننا نريد التعبير عن جودة استقراء معين في حالة فيزيائية معينة . فالتنبؤ بأن السعر

العالمى للنقط سيكون ٤٥ دولار للبرميل الواحد فى السنة القادمة ليس كافياً ، ويجب ألا يشكل حافزاً للدول المستهلكة للشراء أو عدم الشراء .

فالسؤال الأساسى هو ما إذا كان هذا التقدير صحيحاً فى حدود خمسة دولارات زيادة أو نقصاناً . هذا ويجوز الاستقراء الإحصائى لحالة معينة عنصرين أساسيين لاغنى عنهما وهذين العنصرين هما ١ - الاستقراء ٢ - مقياس جودة هذا الاستقراء . والسؤال المطروح هو : أية طريقة أفضل فى الاستقراء ، التقدير أم اختبار الفرضيات ؟ أما الجواب فيعود إلى طبيعة المسألة المدروسة وإلى التفضيل الشخصى . ولا بد من أن نشير إلى أن للتقدير نوعين . فهناك التقدير النقطى والتقدير الجامى . ففى حالة التقدير النقطى يقوم الدارس باستخدام كافة المعلومات المتوفرة من خلال عينة مدروسة للوصول إلى عدد واحد أو نقطة تكون تقديراً للوسيط المراد تقديره . فمثلاً سنجد مستقبلاً أن وسط العينة  $\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$  يمثل تقديراً نقطياً لوسط المجتمع  $\mu$  .

### (٦، ٢) طرق التقدير الكلاسيكية Classical estimation methods

إن تقدير وسيط فى مجتمع إحصائى مدروس يمكن أن يعطى كتقدير نقطى أو كتقدير جامى . يرمز عادة للتقدير النقطى للوسيط  $\hat{\theta}$  فى مجتمع إحصائى بالرمز  $\hat{\theta}$  ولقيمه بالرمز  $\hat{\theta}$  . فمثلاً تمثل القيمة  $\hat{\theta}$  للإحصاء  $X$  والمحسوبة من خلال عينة حجمها  $n$  تقديراً نقطياً للوسيط  $\mu$  فى أى مجتمع إحصائى مدروس .

يسمى الإحصاء المستخدم للحصول على تقدير نقطى بتقدير (estimator) أو بدالة القرار (decision function) . لذلك فإن دالة القرار  $S$  التابعة لقيم العينة تمثل تقديراً للوسيط  $\mu$  ، والتقدير  $S$  هو الإجراء المأخوذ من عينات مختلفة والذي سيؤدى بشكل عام إلى إجراءات مختلفة أو تقديرات .

### تعريف (٦، ١) فضاء القرار Decision space

إن مجموعة كل الإجراءات الممكنة والمأخوذة فى تقدير وسيط ، تدعى الإجراءات أو فضاء القرار .

ولا يمكن تقدير وسيط مجتمع إحصائى بدون ارتكاب خطأ . فلا نتوقع مثلاً أن تكون  $X$  تقديراً للوسيط  $\mu$  بدون أخطاء . ونأمل ألا يكون هذا التقدير بعيداً جداً عن القيمة الحقيقية

للموسيط  $\mu$  . فمن الممكن الحصول على تقدير قريب جداً من الموسيط المجهول  $\mu$  بالنسبة لعينة خاصة وذلك باستخدام متوسط هذه العينة  $\bar{X}$  كتقدير  $\mu$  . لنعتبر مثلاً أن عينة خاصة تتألف من النقاط 4, 8, 9 مأخوذة من مجتمع وسطه  $\mu = 6$  ( ولكن نفترض أنه مجهول ) . نستطيع أن

نقدر  $\mu$  بواسطة العينة  $\bar{x} = 7$  أو  $\bar{x} = 8$  باستخدام متوسط العينة كتقدير لهذا الموسيط . وفي هذه الحالة نلاحظ أن متوسط العينة  $\bar{X}$  يمثل تقديراً أقرب إلى القيمة الحقيقية للموسيط  $\mu$  من التقدير  $\bar{x}$  الممثل لوسط العينة . من ناحية أخرى إذا احتوت عينة على القيم 4, 7, 8, 5 ، فعندئذ نجد أن  $\bar{x} = 7$  و  $\bar{x} = 6.5$  ، ولذلك فإن وسط العينة  $\bar{X}$  يمثل تقديراً أفضل في هذه الحالة . وفي حال عدم معرفة القيمة الحقيقية للموسيط  $\mu$  ، علينا أن نحدد سلفاً أيهما سيكون أفضل  $\bar{X}$  أو  $\bar{x}$  لاستخدامه كتقدير للموسيط  $\mu$  المجهول .

وما يخطر على بال الدارس هو معرفة الخواص التي يجب أن تتمتع بها دالة القرار والتي تؤثر علينا في اختيار تقدير دون آخر يختلف عنه .

لنفرض أن  $\hat{\theta}$  تقديراً تمثل قيمته  $\theta$  تقديراً نقطياً للموسيط المجهول  $\theta$  في مجتمع إحصائي .

**تعريف (٦,٢) تقدير غير متحيز unbiased - estimator**

نسمى الإحصاء  $\hat{\theta}$  تقديراً غير متحيز (unbiased-estimator) للموسيط  $\theta$  المجهول إذا تحقق الشرط التالي :

$$E[\hat{\theta}] = \theta$$

مثال (٦,١)

نلاحظ أن  $\bar{X}$  هو تقدير غير متحيز للموسيط  $\mu$  لأي مجتمع . ذلك لأن :

$$E(\bar{X}) = E\left(\frac{\sum_{i=1}^n X_i}{n}\right) = \frac{\sum_{i=1}^n E(X_i)}{n} = E(X_1) = \mu$$

مثال (٦,٢)

يؤن أن  $S^2$  يمثل تقديراً غير متحيز للتباين  $\sigma^2$  .

الحل

$$\sum_{i=1}^n (X_i - \bar{X})^2 = \sum_{i=1}^n [(X_i - \mu) - (\bar{X} - \mu)]^2$$

نلاحظ أن

$$\begin{aligned} &= \sum_{i=1}^n (X_i - \mu)^2 - 2(\bar{X} - \mu) \sum_{i=1}^n (X_i - \mu) + n(\bar{X} - \mu)^2 \\ &= \sum_{i=1}^n (X_i - \mu)^2 - n(\bar{X} - \mu)^2 \end{aligned}$$

ولكن :

$$\begin{aligned} E(S^2) &= E \left[ \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1} \right] \\ &= \frac{1}{n-1} \left[ \sum_{i=1}^n E(X_i - \mu)^2 - n E(\bar{X} - \mu)^2 \right] \\ &= \frac{1}{n-1} \left( \sum_{i=1}^n \sigma_{X_i}^2 - n \sigma_{\bar{X}}^2 \right) \end{aligned}$$

غير أنه من أجل  $i = 1, 2, \dots, n$  لدينا :

$$\sigma_{X_i}^2 = \sigma^2$$

ومنه :

$$\sigma_{\bar{X}}^2 = \frac{\sigma^2}{n}$$

وأخيراً فإننا نستنتج أن :

$$E(S^2) = \frac{1}{(n-1)} \left( n \sigma^2 - n \frac{\sigma^2}{n} \right) = \sigma^2$$

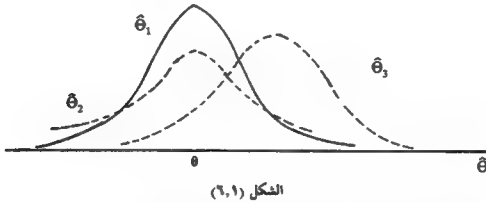
والعلاقة الأخيرة توضح أن  $S^2$  يمثل تقديراً غير متحيزاً للمتوسط  $\sigma^2$  ، ومن جهة ثانية فإن  $S$  هو تقدير متحيز لـ  $\sigma$  بانحياز ضئيل وتافه من أجل عينات كبيرة الحجم .

إذا كان  $\theta_1, \theta_2$  يمثلان تقديرين غير متحيزين لنفس الوسيط  $\theta$  في مجتمع إحصائي ، فإننا نختار التقدير ذي التباين الأقل . لذلك إذا كان  $\sigma_{\theta_1}^2 < \sigma_{\theta_2}^2$  فإننا نقول بأن  $\theta_1$  أكثر فعالية من التقدير  $\theta_2$  .

**تعريف (٦,٣) التقدير غير المتحيز للوسيط Unbiased estimator for median**

نقول بأن التقدير غير المتحيز للوسيط  $\theta$  أكثر فعالية إذا كان تباينه أقل من تباين أى تقدير غير متحيز لنفس الوسيط .

يوضح الشكل (٦,١) توزيعات المعاينة لثلاثة تقديرات  $\theta_1, \theta_2, \theta_3$  للوسيط  $\theta$ . ومن الواضح أن كلا من  $\theta_1, \theta_2$  يمثل تقديرا غير متحيز للوسيط  $\theta$  لأن توزيع كل من هذين التقديرين متناظرا بالنسبة لـ  $\theta$  . كما أن للتقدير غير المتحيز  $\theta_2$  تباين أقل من تباين  $\theta_1$  ، ولذلك فإن  $\theta_1$  أكثر فعالية من  $\theta_2$  ، وعلى هذا الأساس فإننا نختار من بين التقديرات الثلاثة السابقة كتقدير للوسيط  $\theta$  التقدير  $\theta_1$  .



الشكل (٦,١)

لقد أوضحنا أن  $\bar{X}$  يمثل تقديرا غير متحيز للوسط  $\mu$  في أى مجتمع ، ويمكن بسهولة أن نرى أن  $\bar{X}$  متوسط العينة ما هو إلا تقدير غير متحيز للوسط  $\mu$  في المجتمع الطبيعي أيضا . غير أن تباين التقدير  $\bar{X}$  أصغر من تباين التقدير  $\bar{X}$  . وهكذا فإن كلا من القيمتين  $\bar{x}$  و  $\bar{y}$  ستكون في المتوسط مساوية لقيمة المجتمع  $\mu$  ، ولكن  $\bar{X}$  ستكون ( من أجل عينة معينة ) أقرب إلى  $\mu$  .

إن المجال الذى يحتوى فيه التقدير يسمى بمجال التقدير . وهو مجال محدود العرض مركزه التقدير نفسه . ويجب أن يحتوى هذا المجال على القيمة الحقيقية للوسيط المجهول ،

وللحصول على مجال التقدير للوسيط المجهول  $\theta$  علينا أن نشكل مجالاً من الشكل  $\hat{\theta} \pm k$  حيث تتعلق  $\hat{\theta}$  بالعينة الخاصة المختارة ويتحدد العدد  $k$  بواسطة توزيع المعاينة للإحصاء  $\hat{\theta}$ . والآن فإن إدعاءنا بأن التقدير  $\hat{\theta}$  مساو تماماً للوسيط  $\theta$  يعنى أن :

$$\hat{\theta} - k < \theta < \hat{\theta} + k$$

يمكن بواسطة توزيع المعاينة للإحصاء  $\theta$  أن نحدد قيمة  $k$  بحيث يكون الاحتمال  $P[\hat{\theta} - k < \theta < \hat{\theta} + k]$  مساوياً لعدد معين يهم الدارس . فإذا حسبنا مثلاً  $k$  بحيث يكون  $P[\hat{\theta} - k < \theta < \hat{\theta} + k] = 0.95$  . فنعندئذ ستعطينا العينة المسحوبة عشوائياً مجالاً يحتوى على الوسيط المجهول باحتمال قدره 0.95 . يسمى المجال السابق المحسوب من خلال عينة عشوائية بـ  $\pm 95\%$  مجال ثقة . وهذا الاحتمال يعنى أن  $\pm 95\%$  من ثقتنا تؤكد بأن مجالنا هذا سيحتوى على القيمة الحقيقية للوسيط . في المجتمع المدروس ، وكذلك الأمر إذا كان الاحتمال مساوياً لـ  $\pm 0.99$  . فنعندئذ ستكون قيمة  $k$  أكبر ، وبذلك يصبح عرض مجال الثقة أكبر . ونلاحظ أنه كلما عظمت الثقة كلما عرض المجال . وبشكل عام فإن توزيع الإحصاء  $\hat{\theta}$  سيمكننا من حساب  $k$  بحيث يكون :

$$P[\hat{\theta} - k < \theta < \hat{\theta} + k] = 1 - \alpha : 0 < \alpha < 1$$

نسمى المجال السابق المحسوب من خلال عينة خاصة بـ  $100\%(1 - \alpha)$  مجال ثقة . فمثلاً من أجل  $\alpha = 0.05$  نجد  $95\%$  مجال ثقة . ومن أجل  $\alpha = 0.01$  نجد  $99\%$  مجال ثقة . تدعى النسبة  $(1 - \alpha)$  معامل الثقة . وحدود المجال  $\hat{\theta} - k$  و  $\hat{\theta} + k$  بنهايات الثقة أو نهايات الإسناد .

### (٦,٣) تقدير الوسط Estimating the mean

تقدير للوسط  $\mu$  في مجتمع إحصائي بأخذ عادة الإحصاء  $\bar{X}$  . ومن المعلوم أن توزيع المعاينة للإحصاء  $\bar{X}$  يتركز حول الوسط  $\mu$  ، وفي أغلب التطبيقات يكون تباين  $\bar{X}$  أصغر من تباين أى تقدير آخر لـ  $\mu$  . ولذا فإن وسط العينة  $\bar{X}$  سيستخدم كتقدير نقطى للوسط  $\mu$  في أى مجتمع إحصائي ، وإذا تذكرنا بأن  $\sigma_{\bar{X}}^2 = \frac{\sigma^2}{n}$  أدركنا أن العينة الكبيرة ستقدم لنا قيمة لـ  $\bar{X}$  ذات تباين صغير . لذلك فإن من المرجح أن تكون  $\bar{X}$  قريبة جداً من  $\mu$  عندما تكون  $n$  كبيرة . للبحث عن مجال ثقة للوسط  $\mu$  نفرض أن العينة المدروسة مسحوبة من مجتمع طبيعي أو أنها حجم كبير وذلك بغض النظر عن نوع

المجتمع المدروس ، ولنفرض أن تباين المجتمع  $\sigma^2$  معلوم . يمكن أن تشكل مجال ثقة للوسط  $\mu$  بالاعتماد على توزيع المعاينة لـ  $\bar{X}$  . وحسب نظرية النهايات المركزية يمكننا أن نتوقع أن يكون لتوزيع المعاينة  $\bar{X}$  توزيعاً قريباً من التوزيع الطبيعي بالوسط  $\mu_{\bar{X}} = \mu$  والانحراف

المعياري  $\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$  . فإذا فرضنا أن  $z_{\alpha/2}$  هي نقطة من محور السينات تحدد على يمينها مساحة قدرها  $\alpha/2$  لأمكننا كتابة :

$$P(-z_{\alpha/2} < Z < z_{\alpha/2}) = 1 - \alpha$$

حيث إن :

$$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$$

لذلك فإن :

$$P \left( -z_{\alpha/2} < \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} < z_{\alpha/2} \right) = 1 - \alpha$$

لنضرب كل طرف من أطراف المتباينة السابقة بالعدد  $\sigma/\sqrt{n}$  ، ثم لنطرح من جميع الأطراف المقدار  $\bar{X}$  فنجد بعد أن نضرب الأطراف بإشارة ناقص أن :

$$P \left( \bar{X} - z_{\alpha/2} \frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \right) = 1 - \alpha$$

وهكذا نجد أن 100%(1 -  $\alpha$ ) مجال ثقة للوسط  $\mu$  في مجتمع فيه  $\sigma$  معلومة هو المجال :

$$\left( \bar{X} - z_{\alpha/2} \frac{\sigma}{\sqrt{n}} , \bar{X} + z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \right)$$

ملاحظة

إذا كانت العينة مسحوبة من مجتمع غير طبيعي وكان حجمها صغيراً ، فإنه لا

يمكننا التوقع بأن درجة الثقة مضبوطة (أو دقيقة) ، ومع ذلك فإن نظرية العينات تضمن لنا نتيجة جيدة من أجل عينات حجمها  $n \geq 30$  بغض النظر عن شكل المجتمع المدرس .

للبحث عن  $100\%(1 - \alpha)$  مجال ثقة للوسط  $\mu$  نفترض أن  $\sigma$  معلوم . وبما أن هذه الحالة لا تتحقق بشكل دائم ، لذلك يمكن أن نستبدل  $\sigma$  بـ  $S$  شريطة أن تكون  $n \geq 30$  .

### مثال (٣،٦)

يفرض أن أطوال خمسين طالباً ( مسحوبين بشكل عشوائي من مجتمع ما من الطلبة ) قد قدم لنا وسطاً قدره  $\bar{x} = 174.5 \text{ cm}$  وانحرافاً معيارياً قدره  $S = 6.9 \text{ cm}$  . لننشئ  $98\%$  مجال ثقة للوسط  $\mu$  ( أى متوسط طول كل طالب من طلاب المجتمع المدرس ) نلاحظ أن التقدير النقطة للوسط  $\mu$  ( الممثل لمتوسط طول كل طالب ) ما هو إلا  $\bar{x} = 174.5 \text{ cm}$  ، وبما أن حجم العينة المسحوبة من الطلبة هو  $n = 50$  ، لذلك يمكن أن نقرب الانحراف المعياري  $\sigma$  بواسطة  $S = 6.9 \text{ cm}$  . النقطة  $Z$  من محور السينات والتي تحدد على يمينها مساحة قدرها  $0.01$  أو تحدد على يسارها مساحة قدرها  $0.99$  هي النقطة  $Z_{\alpha/2}$  ( أنظر الجدول IV ) . لذلك فإن  $98\%$  مجال الثقة هو المجال :

$$\left( 174.5 - (2.33) \left( \frac{6.9}{\sqrt{50}} \right) , 174.5 + (2.33) \left( \frac{6.9}{\sqrt{50}} \right) \right) = (172.23 , 176.77)$$

وهكذا نجد أن :  $172.23 < \mu < 176.77$

لنبحث عن  $95\%$  مجال ثقة لنفس الوسط  $\mu$  . إن النقطة  $Z$  من محور السينات والتي تحدد على يمينها مساحة قدرها  $0.025$  أو إنها تحدد على يسارها مساحة قدرها  $0.975$  هي النقطة  $Z_{\alpha/2} = 1.9$  ، ولذلك فإن  $95\%$  مجال ثقة للوسط  $\mu$  ما هو إلا المجال :

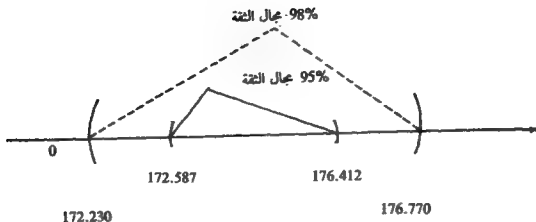
$$\left( 174.5 - (1.96) \left( \frac{6.9}{\sqrt{50}} \right) , 174.5 + (1.96) \left( \frac{6.9}{\sqrt{50}} \right) \right)$$

أو :  $(172.587 , 176.412)$

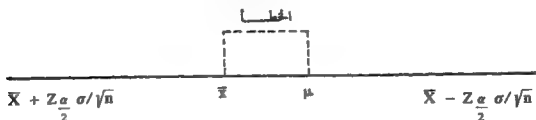
أى إن :  $172.587 < \mu < 176.412$



ويوضح الشكل التالى طولى مجالى الثقة :



من الشكل السابق نستنتج أن مجال الثقة للأطوال يتطلب تقديراً لـ  $\mu$  بدرجة أكبر من الدقة ويزودنا  $100\% (1 - \alpha)$  مجال الثقة بتقدير درجة دقة تقديرنا النقطى ، فإذا كان الوسط المجهول  $\mu$  فعلياً هو مركز المجال فإننا نجد أن التقدير  $\bar{X}$  يقدر  $\mu$  بدون خطأ . وعلى الغالب لا يكون  $\bar{X}$  مساوياً تماماً لـ  $\mu$  ، ولذلك فإن  $\bar{X}$  يقدر  $\mu$  بخطأ . وحجم الخطأ هذا ما هو إلا الفرق بين القيمتين  $\mu$  و  $\bar{X}$  . ويمكننا أن نكون واثقين بنسبة  $(1 - \alpha)$   $100\%$  من أن هذا الفرق سيكون أقل من  $Z_{\alpha/2} \sigma / \sqrt{n}$  ، ويمكننا بسهولة من التحقق مما سبق برسم مخطط لمجال الثقة كما هو موضح على الشكل (٦,٢)



الشكل (٦,٢) يوضح الخطأ المرتكب في تقدير  $\mu$  بـ  $\bar{X}$

### نظرية (٦,١)

إذا كان  $\bar{X}$  تقديراً للوسط  $\mu$  في مجتمع احصائى ما . فإنه يمكننا أن نثق بنسبة قدرها  $100\% (1 - \alpha)$  من الخطأ سيكون أقل من  $Z_{\alpha/2} \sigma / \sqrt{n}$  . ففى المثال (٦,٣) نجد أنه بنسبة قدرها  $98\%$  يمكن أن نثق بأن وسط العينة  $\bar{x} = 174.5 \text{ cm}$  سيختلف عن القيمة

الحقيقية للوسط  $\mu$  ( الممثل لطول كل طالب ) بمقدار أقل من 2.273 cm ، وأن ثنى بنسبة قدرها 95% من أن الفرق سيكون أقل من 1.91 cm .

كثيرا ما نهم في معرفة حجم العينة اللازم للتأكد من أن الخطأ المرتكب في تقدير  $\mu$  سيكون أقل من عدد معين  $e$  بحسب النظرية (٦,١) نرى أن علينا أن نختار  $n$  بحيث يكون :  $Z_{\alpha/2} \sigma / \sqrt{n} = e$  .

$$Z_{\frac{\alpha}{2}} \sigma / \sqrt{n} = e$$

نظرية (٦,٢)

إذا استخدم الإحصاء  $\bar{X}$  كتقدير للوسط  $\mu$  في مجتمع إحصائي ما ، فيمكننا أن ثنى بنسبة  $100\%(1 - \alpha)$  من أن الخطأ المرتكب في التقدير سيكون أقل من عدد معين  $e$  عندما يكون حجم العينة المختارة مساويا لـ :

$$n = \left( \frac{Z_{\alpha/2} \sigma}{e} \right)^2$$

وتكون النظرية (٦,٢) السابقة جاهزة للتطبيق فيما إذا كنا نعرف تباين المجتمع الإحصائي الذي نختار منه عينتنا . وفي الحالة التي تكون فيها  $\sigma$  مجهولة ، فإن بإمكاننا أن نأخذ عينة ( تمهيدية ) ذات حجم  $n \geq 30$  وذلك بغية الحصول على تقدير لـ  $\sigma$  ، وبعد الحصول على هذا التقدير نستخدم النظرية (٦,٢) وبواسطتها نستطيع تقريبا أن نحدد عدد الملاحظات التي نحتاجها للحصول على الدرجة المطلوبة من الدقة .

مثال (٦,٤)

ما هو حجم العينة المطلوب في المثال (٦,٣) إذا أردنا أن يكون الخطأ المرتكب في تقدير  $\mu$  أقل من 2.273 بامثال ثقة قدره 98% ؟

الحل

من الواضح في المثال (٦,٣) المحسوب من عينة تمهيدية حجمها  $n = 50$  أن  $S = 6.9$  cm ، لهذا يمكننا استخدامه عوضا عن  $\sigma$  . وبحسب النظرية (٦,٢) نجد أن :

$$n = \left[ \frac{(2.33)(6.9)}{2.273} \right]^2 = 50.02$$

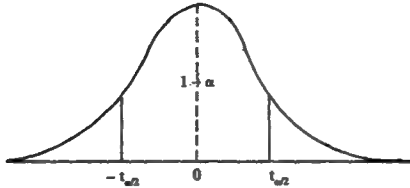
نستنتج مما سبق أن العينة المختارة للتقدير والتي حجمها  $n = 50$  ستقدم لنا تقديراً لـ  $\mu$  لا يتجاوز الخطأ المرتكب فيه 2.273 بثقة قدرها 98% .

غالباً ما نفكر في إيجاد تقدير للوسط  $\mu$  في مجتمع إحصائي تباينه  $\sigma^2$  مجهول . ولكن غالباً ما نضع كلفة العينة والوقت المتوفر ، بالإضافة إلى عوامل أخرى ، حداً لحجم العينة ، بحيث يصبح الحجم الكبير المطلوب غير عملي ، وفي هذه الحالة لا يمكن اللجوء إلى طرق الاستقراء الموافق لعينات حجمها  $n \geq 30$  .

ولقد ذكرنا سابقاً أن للكمية  $\frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$  توزيعاً طبيعياً معيارياً إذا كان المجتمع المدروس طبيعياً ، أو توزيعاً قريباً من التوزيع الطبيعي المعياري إذا كان المجتمع المدروس غير طبيعي ، وذلك طبقاً لنظرية النهاية المركزية بشرط أن يكون حجم العينة  $n$  كبيراً . ولكن في الحالة التي تكون فيها  $n$  صغيرة ، وحيث أنه لا يجوز لنا اللجوء إلى نظرية النهاية المركزية ، فمن الطبيعي وضع  $S$  محسوبة من العينة ( التمهيدية ) بدلاً من  $\sigma$  . وبما أن العينة صغيرة ، فإن هذا يجعلنا نشك في أن تكون  $S$  ممثلة لقيمة تقريبية . وهكذا نفقد كل أساس نظري أو منطقي يسمح لنا بالقول بأن توزيع المتغير الجديد  $\frac{\bar{X} - \mu}{S / \sqrt{n}}$  هو التوزيع الطبيعي المعياري . ومع أن توقع  $S^2$  هو  $\sigma^2 = E(S^2)$  تماماً إلا أن احتمال كون  $S^2$  أقل من  $\sigma^2$  يزيد على النصف كما نرى من توزيع  $S^2$  ] يجب ألا يغيب على ذهن الطالب من أن  $\frac{(n-1)S^2}{\sigma^2}$  له توزيع  $\chi^2_{(n-1)}$  وهذا يحملنا على الاعتقاد بأن منحني الكثافة للمتحويل الجديد  $\frac{\bar{X} - \mu}{S / \sqrt{n}}$  سيكون أكثر انتشاراً .

وعندما تزداد  $n$  ، فإن  $S^2$  تقترب من  $\sigma^2$  ويقترب التوزيع الجديد من التوزيع الطبيعي المعياري . ويرهن في الإحصاء الرياضي أنه إذا كانت  $X_1, X_2, \dots, X_n$  عينة من مجتمع طبيعي وسطه  $\mu$  وتباينه  $\sigma^2$  ، فنعتقد أن يكون للمتغير  $t = \frac{\bar{X} - \mu}{S / \sqrt{n}}$  توزيعاً عينيًا هو التوزيع  $t$  بـ  $(n - 1)$  درجة حرية والذي سبق ذكره في الفصل الخامس .

يوضح الشكل (٦،٣) توزيع المتغير الجديد  $t = \frac{\bar{X} - \mu}{S / \sqrt{n}}$



الشكل (٦،٣)

من هذا الشكل نجد أن :

$$P(-t_{\alpha/2} < T < t_{\alpha/2}) = 1 - \alpha$$

حيث ترمز  $t_{\alpha/2}$  إلى قيمة  $t$  التي تحصر على يمينها مساحة قدرها  $\alpha/2$  ، ومن تناظر الشكل نجد أن مساحة قدرها  $\alpha/2$  مستقع على يسار النقطة  $t = -t_{\alpha/2}$  . وهكذا نجد أن :

$$P\left(-t_{\alpha/2} < \frac{\bar{X} - \mu}{S/\sqrt{n}} < t_{\alpha/2}\right) = 1 - \alpha$$

ويضرب كل طرف من أطراف المتباينة السابقة بـ  $S/\sqrt{n}$  ثم بطرح  $\bar{X}$  وضرب جميع الأطراف بـ 1 - نجد أن :

$$P\left(\bar{X} - t_{\alpha/2} \cdot \frac{S}{\sqrt{n}} < \mu < \bar{X} + t_{\alpha/2} \cdot \frac{S}{\sqrt{n}}\right) = 1 - \alpha.$$

وهكذا نستنتج 100%(1 -  $\alpha$ ) مجال ثقة للوسط  $\mu$  في مجتمع تباينه  $\sigma^2$  مجهول ، وذلك من أجل عينات حجمها  $n < 30$  .

#### مثال (٦،٥)

يزيد وزن نوع معين من الطيور بمقدار 65 غراما خلال الأشهر الثلاثة الأولى من حياتها . وقد أطعمت 12 من هذه الطيور وفقا لنظام تغذية معين خلال الأشهر الثلاثة الأولى وقيست الزيادة في وزن كل منا فوجدت :

55, 62, 54, 58, 65, 64, 60, 60, 62, 59, 67, 62, 61

فالمطلوب البحث عن 95% مجال ثقة للوسط  $\mu$  الممثل لمتوسط الزيادة في الوزن خلال الأشهر الثلاثة الأولى .

الحل

نلاحظ أن  $n = 12$  وأن  $n - 1 = 11$  .

ثم إن :

$$\bar{x}_0 = \frac{\sum_{i=1}^{12} x_i}{12} = 60.75$$

$$S^2 = \frac{\sum_{i=1}^{12} (x_i - \bar{x}_0)^2}{(n - 1)}$$

$$S = 3.8406$$

نلاحظ أن :

$$\bar{X} = 60.75$$

ومجال الثقة المطلوب هو :

$$\left( \bar{X} - \frac{t_{\alpha/2} \cdot S}{\sqrt{n}}, \bar{X} + \frac{t_{\alpha/2} \cdot S}{\sqrt{n}} \right)$$

ومن الجدول VI وبفرض أن  $n - 1 = 11$  نجد أن :

$$t_{\alpha/2}(n - 1) = t_{0.025}(11) = 2.20$$

وهكذا نجد أن مجال الثقة هو :

$$60.75 - (2.20) \frac{(3.8406)}{\sqrt{12}}, 60.75 + (2.20) \frac{(3.8406)}{\sqrt{12}}$$

وأخيراً فإن 95% مجال ثقة للوسط المجهول  $\mu$  هو المجال :  
(58.3108, 63.1891)

#### (٦،٤) تقدير فرق وسطين Estimating the difference between two means

توازي مسألة تقدير الفرق بين وسطين في أهميتها مسألة تقدير وسط مجتمع إحصائي . فقد يرغب أحدنا في مقارنة طريقتين في التعليم ، فنقسم الطلبة عشوائياً إلى زميرتين ، ونخضع كلا منهما لطريقة تعليم معينة ، ثم نقوم بالاستقراء حول الفرق بين تحصيل الطلبة في كل من الزمرتين وذلك بواسطة مقياس معين . أو قد نرغب في مقارنة معدل الإنتاج في شركة للصناعة النسيجية عند استخدام المواد الأولية الواردة من مولين B, A ، فنأخذ عينة من الإنتاج اليومي في كل حالة ونستخدم المعلومات التي تقدمها هاتين العينتين للقيام باستقراء يتعلق بالفرق بين معدل الإنتاج .

فيفرض أن المجتمع الإحصائي الأول بالوسط  $\mu_1$  والتباين  $\sigma_1^2$  ، والثاني بالوسط  $\mu_2$  والتباين  $\sigma_2^2$  ، وأتينا مسحا عينة عشوائية ذات حجم  $n_1$  من المجتمع الأول ، وعينة أخرى ذات حجم  $n_2$  من المجتمع الثاني ولنفرض أن العينتين قد سحبنا بحيث تكون كلا منهما مستقلة عن الأخرى . بعد ذلك نقوم بحساب تقديرات لوسطاء المجتمعين وهي  $\bar{X}_1$  ،  $S_1^2$  من العينة الأولى و  $\bar{X}_2$  ،  $S_2^2$  من العينة الثانية ، والتقدير النقطي للفرق بين وسطى المجتمعين  $\mu_1 - \mu_2$  هو  $\bar{X}_1 - \bar{X}_2$  . وكما نعلم من خواص التوقع أن :

$$E(\bar{X}_1 - \bar{X}_2) = E(\bar{X}_1) + E(-\bar{X}_2) = E(\bar{X}_1) - E(\bar{X}_2) = \mu_1 - \mu_2$$

وبما أن العينتين مستقلتان لذلك فإنه يمكننا كتابة :

$$\sigma_{\bar{X}_1 - \bar{X}_2}^2 = \sigma_{\bar{X}_1}^2 + \sigma_{\bar{X}_2}^2 = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$$

ومنه :

$$\sigma_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

وبالإضافة لذلك إذا كانت كل من  $n_1, n_2$  أكبر من تساوى 30 ، فعندئذ حسب النظرية (٥,١٥) سيكون للمتغير :

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

توزيعاً قريباً من التوزيع الطبيعي المعياري .

أى إنه يمكن أن نخزم باحتمال قدره  $(1 - \alpha)$  من أن هذا المتغير الطبيعي المعياري سيقع ضمن العددين  $Z_{\alpha/2}$  و  $-Z_{\alpha/2}$  أى أن :

$$P[-Z_{\alpha/2} < Z < Z_{\alpha/2}] = 1 - \alpha$$

وبتعويض قيمة  $Z$  في المعادلة السابقة نجد أن :

$$P \left[ -Z_{\alpha/2} < \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} < Z_{\alpha/2} \right] = 1 - \alpha$$

وبضرب كل حد من حدود المتباينة السابقة بالعدد الموجب :

$$\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

ثم نطرح العدد  $(\bar{X}_1 - \bar{X}_2)$  من كل حد من حدودها ، وأخيراً بضرب جميع حدود المتباينة بـ  $-1$  نجد أن :

$$P \left[ (\bar{X}_1 - \bar{X}_2) - Z_{\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} < \mu_2 - \mu_2 < (\bar{X}_1 - \bar{X}_2) + Z_{\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \right] = 1 - \alpha$$

وهكذا نصل إلى التعريف التالى :

تعريف مجال ثقة فرق وسطين  $\mu_1 - \mu_2$

إن  $100\%(1 - \alpha)$  مجال ثقة فرق وسطين  $\mu_1 - \mu_2$  هو المجال :

$$\bar{X}_1 - \bar{X}_2 - Z_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} < \mu_1 - \mu_2 < (\bar{X}_1 - \bar{X}_2) + Z_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

وذلك بفرض أن تباينى المجتمعين المدروسين  $\sigma_1^2, \sigma_2^2$  معلومان ، وأن  $Z_{\alpha/2}$  هى النقطة من المحور السينى تحت المنحنى الطبيعي المعيارى التى تحدد تحت المنحنى ، وعلى يمينها مساحة قدرها  $\alpha/2$  .

ملاحظة :

كما ذكرنا سابقا إذا كان  $\sigma_1^2, \sigma_2^2$  مجهولين وكانت العينات المختارة كبيرة بشكل كاف ( أى أن  $n_1 \geq 30$  و  $n_2 \geq 30$  ) فعندئذ يمكن استبدال  $\sigma_1^2, \sigma_2^2$  بـ  $S_1^2, S_2^2$  على التالى .

مثال (٦,٦)

لدى مقارنة نوعين من إطارات السيارات باختبار عملي أخذنا من كل نوع عينة حجمها  $n = 100$  إطاراً ، ثم سجلنا عدد الأميال التى خدّمها كل إطار حتى اهترائه وفقاً لمقاييس محددة سلفاً . فإذا كانت نتائج الاختبار بالأميال كالتالى :

$$\begin{aligned} \bar{X}_1 &= 26400 & S_1^2 &= 1440000 \\ \bar{X}_2 &= 25100 & S_2^2 &= 1960000 \end{aligned}$$

فما هو تقدير الفرق بين وسطي العمر في النوعين ، ثم ما ماهو حدود الخطأ في التقدير ؟



## الحل

نلاحظ أن التقدير التقطى لـ  $(\mu_1 - \mu_2)$  حيث يمثل  $\mu_1$  العمر الوسطى لإطار من النوع الأول ،  $\mu_2$  العمر الوسطى للإطار من النوع الثاني [ هو  $\bar{X}_1 - \bar{X}_2$  أى :

$$\bar{X}_1 - \bar{X}_2 = 26400 - 25100 = 1300$$

وكذلك نجد أن :

$$\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}} = 184$$

ونلاحظ أن حدود الخطأ هو حوالى  $2\sigma_{\bar{x}_2 - \bar{x}_1}^2 = 368$  ، وهكذا نجد أن النوع الأول متفوق على النوع الثاني .

## مثال (٦,٧)

نبحث عن 0.99 مجال ثقة للفرق  $\mu_1 - \mu_2$  فى المثال (٦,٦) .

## الحل

نعلم أن  $Z_{\alpha/2} = 2.58$  ومنه فمجال الثقة المطلوب هو

$$(\bar{x}_1 - \bar{x}_2) - 2.58 \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} , (\bar{x}_1 - \bar{x}_2) + 2.58 \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

وباستخدام نتائج التمرين (٦,٦) نجد أن 0.99 مجال ثقة لفرق الوسطين  $\mu_1$  ,  $\mu_2$  هو المجال :

$$(1300 - 2.58 (184) , 1300 + 2.58 (184))$$

أو المجال :

$$(825.28 , 1774.72)$$

أى أن الحد الأدنى للثقة هو 825 ، والحد الأعلى للثقة هو 1775 ، أى أنه يمكن تقدير الفرق بين وسطي العمرين بأنه يقع بين هذين الحدين .

## ملاحظة

إذا كانت العينات المختارة من المجتمعين صغيرة الحجم ، فعندئذ نلجأ إلى التوزيع ؛ لإيجاد مجالات ثقة وهذا الأمر صحيح وشرعى عندما تكون المجتمعات المدروسة لها تقريبا توزيعات طبيعية .

نفرض الآن أن  $\sigma_1^2, \sigma_2^2$  مجهولين وأن  $n_1, n_2$  صغيرين ، وكلا منهما أقل من 30 . إذا كان  $\sigma_1^2 = \sigma_2^2 = \sigma^2$  فإننا نحصل على متغير عشوائى طبيعى معيارى من الشكل :

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\sigma^2 \left[ \frac{1}{n_1} + \frac{1}{n_2} \right]}}$$

وبحسب النظرية (٥، ١٦) يكون للمتغيرين  $(n_1 - 1)S_1^2/\sigma^2, (n_2 - 1)S_2^2/\sigma^2$  توزيعين من نوع  $\chi^2$  ( كاي - مربع ) بعدد من درجات الحرية مساو على الترتيب  $(n_1 - 1), (n_2 - 1)$  . وهذين المتغيرين مستقلين أيضا ( لأن العينات المختارة مستقلة ) . لذلك فإن مجموعهما :

$$V = \frac{(n_1 - 1) S_1^2}{\sigma^2} + \frac{(n_2 - 1)S_2^2}{\sigma^2} \quad . \quad = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{\sigma^2}$$

له توزيع من نوع  $\chi^2$  ( كاي - مربع ) بعدد من درجات الحرية مساو لـ  $(n_1 + n_2 - 2)$  . وبتعويض  $Z, V$  فى النظرية (٥، ١٧) فإننا نحصل على الإحصاء التالى :

$$T = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\sigma^2 \left[ \frac{1}{n_1} + \frac{1}{n_2} \right]}} \sqrt{\frac{(n_1 - 1) S_1^2 + (n_2 - 1) S_2^2}{\sigma^2 (n_1 + n_2 - 2)}}$$

ويتوزع هذا المتغير وفقا للتوزيع ؛ بعدد من درجات الحرية مساو لـ  $(n_1 + n_2 - 2)$  ، وأن التقدير التقطى للتباين المجهول  $\sigma^2$  يمكن الحصول عليه بواسطة المجموع :

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

وبتبديل  $S_p^2$  في عبارة الإحصاء  $T$  نجد أن :

$$T = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

وباستخدام الإحصاء السابق نجد أن :

$$P [-t_{\alpha/2} < T < t_{\alpha/2}] = 1 - \alpha$$

حيث تمثل  $t_{\alpha/2}$  القيمة  $t$  الواقعة على محور السينات أسفل منحني التوزيع  $t$  بـ  $(n_1 + n_2 - 2)$  درجة من الحرية ، والتي تحدد على يمينها مساحة قدرها  $\alpha/2$  . وبتبديل قيمة  $T$  في المعادلة السابقة نجد أن :

$$P \left[ -t_{\alpha/2} < \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} < t_{\alpha/2} \right] = 1 - \alpha$$

أو العلاقة :

$$P \left[ (\bar{X}_1 - \bar{X}_2) - t_{\alpha/2} S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} < \mu_1 - \mu_2 < (\bar{X}_1 - \bar{X}_2) + t_{\alpha/2} S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \right] = 1 - \alpha$$

فمن أجل عيتين عشوائيتين مستقلتين ذات حجمين  $n_1, n_2$  مختارتين بشكل عشوائي ، من مجتمعين طبيعيين نجد أن فرق وسطى العيتين  $\bar{X}_1 - \bar{X}_2$  يمثل تقديرا نقطيا لفرق الوسطين  $\mu_1 - \mu_2$  ، وأن 100%  $(1 - \alpha)$  مجال الثقة لهذا التقدير هو المجال :

$$\left( (\bar{X}_1 - \bar{X}_2) - t_{\alpha/2} S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}, (\bar{X}_1 - \bar{X}_2) + t_{\alpha/2} S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \right)$$

حيث يمثل  $S_p$  الانحراف المعياري المشترك للمجتمعين المدروسين .

مثال (٦,٨)

استخدم في عملية كيميائية عاملان مساعدان لمقارنة تأثيرهما على قدرة عملية

التفاعل . وقد تم تحضير اثني عشر مزيجاً باستخدام العامل المساعد الأول ، فأعطت هذه الأمزجة وسطاً قدره  $\bar{X}_1 = 85$  وانحرافاً معيارياً قدره  $S_1 = 4$  ، كما تم تحضير عشرة أمزجة باستخدام العامل المساعد الثاني ، فأعطت وسطاً قدره  $\bar{X}_2 = 81$  وانحرافاً معيارياً قدره  $S_2 = 5$  .

فالمطلوب البحث عن 90% مجال ثقة للفرق بين وسطى المجتمعين ، وذلك بفرض أن للمجتمعين توزيعاً قريباً من التوزيع الطبيعي بنفس التباين  $\sigma^2$  . بفرض أن  $\mu_1$  ،  $\mu_2$  يمثلان على الترتيب وسطى المجتمعين اللذين أخذ منهما العاملان المساعدان الأول والثاني . فنعتمد أن يكون المطلوب البحث عن 90% مجال ثقة للفرق  $\mu_1 - \mu_2$  .  
نأخذ  $\bar{X}_1 - \bar{X}_2 = 85 - 81 = 4$  كتقدير نقطي للفرق فنلاحظ أن :

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

$$= \frac{(11)(16) + (9)(25)}{12 + 10 - 2} = 20.05$$

وبأخذ الجذر التربيعي للطرفين نجد أن  $S_p = 4.478$  . والملاحظ أيضاً أن  $1 - \alpha = 0.90$  ، إذا  $\alpha = 0.1$  . ومن الجدول  $t$  نجد أن  $t_{0.05} = 1.725$  من أجل عدد من درجات الحرية  $\nu = n_1 + n_2 - 2 = 20$  .

لذلك بالتعويض في العلاقة :

$$(\bar{x}_1 - \bar{x}_2) - t_{\alpha/2} S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} < \mu_1 - \mu_2 < (\bar{x}_1 - \bar{x}_2) + t_{\alpha/2} S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

نجد أن 90% مجال الثقة هو التالي :

$$4 - (1.725)(4.478) \sqrt{\frac{1}{12} + \frac{1}{10}} < \mu_1 - \mu_2 < 4 + (1.725)(4.478) \sqrt{\frac{1}{12} + \frac{1}{10}}$$

أى مجال :

$$0.69 < \mu_1 - \mu_2 < 7.31$$

وحيث أن نهايات الثقة موجبة ، لذلك فإننا نستنتج أن العامل المساعد الأول متفوق على العامل المساعد الثاني .

لنفتش عن مجال ثقة للفرق  $\mu_1 - \mu_2$  من أجل عينات صغيرة ، وذلك بفرض أن تباينات المجتمعات المدروسة مجهولة ، وغير متساوية ، وحجوم العينات المدروسة مختلفة أيضا . في هذه الحالة نلاحظ أن للإحصاء .

$$T^* = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1, \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

توزيعاً قريباً من التوزيع  $t$  بعدد من درجات الحرية قدره :

$$= \frac{[(S_1^2 / n_1) + (S_2^2 / n_2)]^2}{[(S_1^2 / n_1)^2 / (n_1 - 1)] + [(S_2^2 / n_2)^2 / (n_2 - 1)]}$$

وحيث إن  $v$  عدداً نادراً ، لذلك نقربه إلى أقرب عدد تام . وباستخدام الإحصاء  $T^*$  نجد أن :

$$P(-t_{\alpha/2} < T < t_{\alpha/2}) \approx 1 - \alpha$$

حيث يمثل  $t_{\alpha/2}$  قيمة المتغير  $t$  بعدد من درجات الحرية مساو لـ ، والتي تحدد على يمينها مساحة قدرها  $\alpha/2$  ، ويتمويض قيمة  $T$  نجد أن :

$$(\bar{X}_1 - \bar{X}_2) - t_{\alpha/2} \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}} < \mu_1 - \mu_2 < (\bar{X}_1 - \bar{X}_2) + t_{\alpha/2} \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$$

حيث يمثل  $\bar{X}_1$  ،  $\bar{X}_2$  وسطى العيتين الصغيرتين المستقلتين ذات الحجمين  $n_1$  ،  $n_2$  ، والمأخوذتين من مجتمعين هما تقريباً توزيعاً طبيعياً . وكذلك فإن  $S_1^2$  ،  $S_2^2$  يمثلان أيضاً التباينين العينين لهاتين العيتين .

## مثال (٦,٩)

لدى الرجوع إلى سجلات الأرصاد الجوية في المملكة العربية السعودية ، وجد أن وسط هطول الأمطار التي هطلت خلال الخمسة عشر عاما الماضية على منطقة القصيم في شهر نيسان هو 4.93 سنتيمتر ، وأن انحرافها المعياري هو 1.14 سنتيمتر ، كما وجد أن وسط هطول الأمطار التي سقطت في منطقة مجاورة للقصيم في نفس الشهر خلال العشرة أعوام السابقة من نفس الشهر هو 2.64 سنتيمتر ، والانحراف المعياري لها هو 0.66 سنتيمتر . لننشئ 95% مجال ثقة لفرق الوسطين الحقيقيين لهطول الأمطار في هاتين المنطقتين ، وذلك بفرض أن الملاحظات السابقة أخذت من مجتمعات طبيعية بتباينات مختلفة . من أجل منطقة القصيم لدينا  $\bar{X}_1 = 4.93$  ،  $S_1 = 1.14$  ،  $n_1 = 15$  ، ومن أجل المنطقة المجاورة للقصيم لدينا  $\bar{X}_2 = 2.64$  ،  $S_2 = 0.66$  ،  $n_2 = 10$  . لنفرض أن  $\mu_1$  هو الوسط الحقيقي لهطول الأمطار في القصيم خلال عام و  $\mu_2$  في المنطقة المجاورة ، وحيث أن تباينى المجتمعين المدروسين الطبيعيين مختلفان وحجمي العييتين مختلفان أيضا ، لذلك يمكننا أن نجد 95% مجال ثقة تقريبا بالاعتماد على التوزيع t بعدد من درجات الحرية :

$$v = \frac{[S_1^2/n_1 + S_2^2/n_2]^2}{\{[S_1^2/n_1]^2/n_1 - 1\} + \{[S_2^2/n_2]^2/n_2 - 1\}}$$

$$= \frac{[(1.14)^2/15 + (0.66)^2/10]^2}{\{[(1.14)^2/15]^2/14\} + \{(0.66)^2/10\}^2/9}$$

ومنه :

$$v = 22.7 \approx 23$$

كما نلاحظ أن التقدير النقطة للفرق  $\mu_1 - \mu_2$  هو  $\bar{X}_1 - \bar{X}_2 = 4.93 - 2.64 = 2.29$  ، وحيث أن  $1 - \alpha = 0.95$  ، لذلك فإن  $\alpha = 0.05$  . من الجدول V نجد أن  $t_{\alpha/2} = t_{0.025} = 2.069$  . وبالتعويض في العلاقة :

$$(\bar{x}_1 - \bar{x}_2) - t_{\alpha/2} \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}} < \mu_1 - \mu_2 < (\bar{x}_1 - \bar{x}_2) + t_{\alpha/2} \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$$

نحصل على المجال الثقة :

$$2.02 < \mu_1 - \mu_2 < 2.56$$

لذلك فإن 95% من ثقتنا تؤكد أن المجال (2.02, 2.56) يحوى الفرق الحقيقى المتوسطى  
مطول الأمطار فى هاتين المنطقتين .

### (٦,٥) تقدير P في المجتمع الحداني Estimating P in binomial proportions

إن أفضل تقدير نقطى لـ P في المجتمع الحداني ، هو وسط عدد النجاحات خلال  
n تكراراً . أى أن التقدير P هو :

$$\hat{p} = \frac{X}{n}$$

حيث يمثل X عدد النجاحات و n عدد التكرارات ، ونقصد بعبارة ( أفضل تقدير ) أن  
التقدير  $\hat{p}$  غير متحيز وله تباين أصغر من تباين أى تقدير غير متحيز لـ  $\hat{p}$  .

وفقا لنظرية النهايات المركزية ، فإن التوزيع التقريبي لـ X هو التوزيع الطبيعي  
بالوسط np والتباين np.q وذلك من أجل قيمة كبيرة لـ n ، وبما أن 1/n يمثل عدداً ثابتاً  
لذلك ، حسب خاصية التوزيع الطبيعي المذكورة في الفصل الخامس ، ويكون التوزيع  
التقريبى لـ  $\hat{p} = \frac{1}{n}(X)$  هو أيضا التوزيع الطبيعي بالوسط :

$$E(\hat{p}) = E\left(\frac{X}{n}\right) = \frac{1}{n} \cdot E(X) = \frac{1}{n} \cdot np = p$$

والتباين :

$$\sigma_{\hat{p}}^2 = \sigma_{\frac{X}{n}}^2 = \frac{1}{n^2} \sigma_X^2 = \frac{npq}{n^2} = \frac{pq}{n}$$

ويمكننا التأكد من أن :

$$P\left(-Z_{\alpha/2} < \frac{\hat{p} - p}{\sqrt{pq/n}} < Z_{\alpha/2}\right) = 1 - \alpha$$

حيث تمثل  $Z_{\alpha/2}$  قيمة على المحور السينى تحت منحنى دالة الكثافة للمتغير الطبيعي  
المعياري ، والتي تحدد على يمينها مساحة قدرها  $\alpha/2$  . هذا ويمكن كتابة العلاقة الأخيرة  
على النحو التالى :

$$P\left(\hat{p} - Z_{\alpha/2} \sqrt{\frac{pq}{n}} < p < \hat{p} + Z_{\alpha/2} \sqrt{\frac{pq}{n}}\right) = 1 - \alpha$$

والصعوبة التي تواجهنا هنا هي في حساب  $\sigma_n \frac{Pq}{n}$  التي تعتمد على  $P$  وعلى  $q = 1 - p$  وهما مجهولان . فإذا وضعنا  $P, q$  بدلا عن  $q, P$  في عبارة الانحراف المعياري  $\sqrt{\frac{Pq}{n}}$  فإن الخطأ المرتكب سيكون صغيراً جداً من أجل قيمة كبيرة لـ  $n$  . وفي الحقيقة يتغير الانحراف المعياري ببطء شديد عندما تتحول  $P$  . ويمكن ملاحظة ذلك بوضوح في الجدول التالي :

$P$	$\sqrt{Pq}$
0.5	0.50
0.4	0.49
0.3	0.46
2.0	0.40
1.0	0.30

ويجدر الانتباه إلى أن تغير  $\sqrt{Pq}$  طفيف جداً من أجل قيم  $P$  القريبة من 0.5 . لذلك فإنه بتبديل التقدير النقطي  $\hat{P} = \frac{X}{n}$  بدلا من  $p$  في داخل الجذر نجد أن :

$$P \left( \hat{P} - Z_{\alpha/2} \sqrt{\frac{\hat{P}q}{n}} < P < \hat{P} + Z_{\alpha/2} \sqrt{\frac{\hat{P}q}{n}} \right) \approx 1 - \alpha$$

وهكذا نجد أننا سحبنا عينة ذات حجم  $n$  من مجتمع حداني فيه  $\hat{P}$  مجهول وحسبنا  $\hat{P} = \frac{X}{n}$  من خلال هذه العينة فنعدّد تحت المجال :

$$\hat{P} - Z_{\alpha/2} \sqrt{\frac{\hat{P}q}{n}} < P < \hat{P} + Z_{\alpha/2} \sqrt{\frac{\hat{P}q}{n}}$$

مثلاً لـ 100% (1 -  $\alpha$ ) مجال ثقة للوسيط  $P$  .

#### ملاحظة

إن المجال السابق هو 100% (1 -  $\alpha$ ) مجال ثقة لـ  $p$  محسوباً من عينة عشوائية حجمها



$n \geq 30$  مختارة من مجتمع حداني .

مثال (٦,١٠)

أعطيت نتائج استطلاع للرأى حول انتخاب السيد X في منطقة معينة من بين مائة ناخب ، أن  $x = 59$  ناخبا يؤيدون ترشيح المرشح المذكور . لنبحث عن تقدير نسبة الناخبين في المنطقة الذين يفضلون هذا الناخب ولنضع حدوداً لخطأ هذا التقدير .

نلاحظ أن التقدير النقطي لـ P هو  $\hat{p} = \frac{59}{100} = 0.59$  . وحدود الخطأ المرتكب في هذا التقدير هي حوالى :

$$2 \sqrt{\frac{\hat{p} \hat{q}}{n}} = 2 \sqrt{\frac{(0.59)(0.4)}{100}} = 0.096$$

كما أن 95% مجال ثقة لـ p هو المجال :

$$\left( \hat{p} - Z_{\alpha/2} \sqrt{\frac{\hat{p} \hat{q}}{n}} , \hat{p} + Z_{\alpha/2} \sqrt{\frac{\hat{p} \hat{q}}{n}} \right)$$

وبالتعويض نجد أن :

$$0.59 - (1.96)(0.096) , 0.59 + (1.96)(0.096)$$

أى المجال :

$$0.4018 < p < 0.7781$$

مثال (٦,١١)

لدى دراسة النسبة الفعلية لعدد التلفزيونات الملونة الموجودة في مدينة جدة بالمملكة العربية السعودية اخترنا عينة عشوائية مؤلفة من 450 عائلة تقتنى أجهزة تلفزيونية ، فوجدنا أن عدد الأجهزة الملونة بينها هو  $x = 300$  . لنفتش عن 90% مجال ثقة لنسبة التلفزيونات الملونة المستخدمة في هذه العملية .

إن التقدير النقطي للنسبة p هو التقدير  $\hat{p} = \frac{300}{450} = 0.666$  باستخدام الجدول IV

نجد أن :

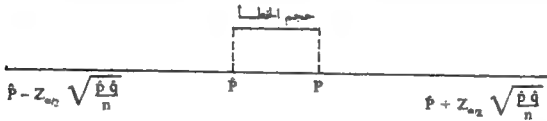
$Z_{0.05} = 1.645$  ، ومن المعلوم أن 90% مجال ثقة للنسبة  $p$  هو المجال :

$$\bar{p} - Z_{\alpha/2} \sqrt{\frac{\bar{p}q}{n}} < p < \bar{p} + Z_{\alpha/2} \sqrt{\frac{\bar{p}q}{n}}$$

وبالتعويض نجد أن :

$$0.63 < p < 0.70$$

نلاحظ أنه إذا كانت القيمة  $p$  هي مركز الـ 100% (1 -  $\alpha$ ) مجال الثقة ، عندئذ يقدر  $\bar{p}$  النسبة  $p$  دون حدوث أى خطأ . وفي أغلب الأحوال لا تتساوى  $\bar{p}$  مع  $p$  ، وهذا يعنى أن هناك خطأ مركباً في التقدير ، وحجم هذا الخطأ هو الفرق بين القيمتين  $p$  و  $\bar{p}$  . ويمكننا بثقة قدرها 100% (1 -  $\alpha$ ) من أن نقول بأن هذا الخطأ أقل من  $Z_{\alpha/2} \sqrt{\frac{\bar{p}q}{n}}$  . ويمكن بسهولة أن نرى ذلك في المخطط التالى :



### نظرية (٦,٣)

إذا كان  $p$  تقديراً للنسبة  $p$  في مجتمع حدائى . فيمكن أن نثق بنسبة قدرها 100% (1 -  $\alpha$ ) من أن الخطأ المركب في هذا التقدير سيكون أقل من  $Z_{\alpha/2} \sqrt{\frac{\bar{p}q}{n}}$  . ففي المثال (٦,١١) كانت نسبة ثقتنا 90% من أن  $p = 0.666$  تختلف عن النسبة الحقيقية  $p$  بمقدار أقل من 0.035 (المثلة لنصف طول مجال الثقة) .

لنحدد الآن الحجم الضرورى للعينة المختارة ليكون  $p$  تقديراً للنسبة  $p$  بخطأ لا يتجاوز العدد  $e = Z_{\alpha/2} \sqrt{\bar{p}q/n}$  مثلاً .

### نظرية (٦,٤)

إذا كان  $\bar{p}$  تقديراً لـ  $p$  فعندئذ يمكن أن نثق بنسبة 100% (1 -  $\alpha$ ) من أن الخطأ المركب في

عملية التقدير هذه لن يتجاوز العدد  $e$  إذا كان حجم العينة :

$$n = Z_{\alpha/2}^2 \frac{p \cdot q}{e^2}$$

مثال (٦,١٢)

ما هو حجم العينة المطلوبة للتأكد من أن الخطأ المرتكب في تقدير  $P$  في المسألة (٦,١١) لن يبلغ 0.03 بثقة قدرها 90% ؟

من الفرض لدينا  $p = 0.666$  ، واعتاداً على النظرية (٦,٤) نجد أن :

$$n = \frac{Z_{\alpha/2}^2 \cdot p \cdot q}{e^2} = \frac{(1.645)^2 (0.666) (0.334)}{(0.03)^2} = 668.8$$

أي أن  $n = 669$  جهاز تلفزيوني . وهذا يعني أننا إذا أسسنا تقديراً للنسبة  $P$  من خلال عينة حجمها  $n = 669$  ، فإنه يمكننا أن نقول بـ 90% من أن تقديرنا  $\hat{p}$  لن يختلف عن القيمة الحقيقية  $P$  بأكثر من 0.03 .

(٦,٦) تقدير الفرق بين نسبتي مجتمعين حدانيين

Estimating the difference between two proportions

نفرض أننا أمام مجتمعين حدانيين بالنسبة  $P_1, P_2$  . لإيجاد تقدير لفرق هاتين النسبتين  $P_1 - P_2$  مثلاً نختار عينة من المجتمع الأول حجمها  $n_1$  ، وعينة ثانية من المجتمع الثاني حجمها  $n_2$  ، وبصورة تكون معها العيتان مستقلتين ، ثم نحسب من خلال هاتين العيتين نسبة النجاح في كل عينة ، أي نقوم بحساب  $P_1, P_2$  ، فيكون الفرق  $P_1 - P_2$  تقديراً نقطياً وغير متحيز للفرق  $P_1 - P_2$  ذلك لأن :

$$E(\hat{P}_1 - \hat{P}_2) = P_1 - P_2$$

أما انحرافه المعياري فيساوي :

$$\sigma_{\hat{P}_1 - \hat{P}_2} = \sqrt{\frac{P_1 q_1}{n_1} + \frac{P_2 q_2}{n_2}}$$

ومن المعلوم أن لكل من  $\hat{P}_1, \hat{P}_2$  توزيعا قريبا من التوزيع الطبيعي بالوسط  $P_1 - P_2$  على الترتيب والتباين  $\frac{P_1 Q_1}{n_1}$  و  $\frac{P_2 Q_2}{n_2}$  . على الترتيب أيضا . وحيث أن العيتين المسحوبتين مستقلتان ، لذلك فإن المتغيرين  $\hat{P}_1, \hat{P}_2$  مستقلان أيضا .

وحسب النظرية (٥,١١) نستنتج أن للمتغير  $\hat{P}_1, \hat{P}_2$  توزيعا قريبا من التوزيع الطبيعي بالوسط  $P_1 - P_2$  والتباين  $\frac{P_1 Q_1}{n_1} + \frac{P_2 Q_2}{n_2}$  لذلك فإن باستطاعتنا التأكد من أن :

$$P \left[ -Z_{\alpha/2} < \frac{(\hat{P}_1 - \hat{P}_2) - (P_1 - P_2)}{\sqrt{\frac{P_1 Q_1}{n_1} + \frac{P_2 Q_2}{n_2}}} < Z_{\alpha/2} \right] = 1 - \alpha$$

حيث تمثل  $Z_{\alpha/2}$  قيمة المتغير العشوائى الطبيعي المعيارى والتي تحدد على يمينها ونحت منحنى الكثافة مساحة قدرها  $\alpha/2$  .

وبإجراءات بعض التغيرات فى أطراف المتباينة السابقة نجد أن :

$$P \left( \hat{P}_1 - \hat{P}_2 - Z_{\alpha/2} \sqrt{\frac{P_1 Q_1}{n_1} + \frac{P_2 Q_2}{n_2}} < P_1 - P_2 < \hat{P}_1 - \hat{P}_2 + Z_{\alpha/2} \sqrt{\frac{P_1 Q_1}{n_1} + \frac{P_2 Q_2}{n_2}} \right) = 1 - \alpha$$

فإذا كان حجم العينتين  $n_1, n_2$  كبيران . فإن بإمكاننا تبديل  $P_1, P_2$  تحت إشارة الجذر بالتقديرين :  $\hat{P}_1 = \frac{x_1}{n_1}$  ،  $\hat{P}_2 = \frac{x_2}{n_2}$  على الترتيب لذلك فإن :

$$P \left( \hat{P}_1 - \hat{P}_2 - Z_{\alpha/2} \sqrt{\frac{\hat{P}_1 \hat{Q}_1}{n_1} + \frac{\hat{P}_2 \hat{Q}_2}{n_2}} < P_1 - P_2 < \hat{P}_1 - \hat{P}_2 + Z_{\alpha/2} \sqrt{\frac{\hat{P}_1 \hat{Q}_1}{n_1} + \frac{\hat{P}_2 \hat{Q}_2}{n_2}} \right)$$

وهكذا نجد أن  $100\%(1 - \alpha)$  مجال ثقة للفرق  $P_1 - P_2$  هو المجال :

$$\left( \hat{P}_1 - \hat{P}_2 - Z_{\alpha/2} \sqrt{\frac{\hat{P}_1 \hat{Q}_1}{n_1} + \frac{\hat{P}_2 \hat{Q}_2}{n_2}} , \hat{P}_1 - \hat{P}_2 + Z_{\alpha/2} \cdot \sqrt{\frac{\hat{P}_1 \hat{Q}_1}{n_1} + \frac{\hat{P}_2 \hat{Q}_2}{n_2}} \right)$$

## مثال (٦,١٣)

لمقارنة فاعلية نوعين A, B من المحاليل المضادة للبعوض ، استخدمت غرفتان من نفس الحجم تحتوي كلا منهما على 500 بعوضة ، وعولجت إحداهما بكمية معينة من المحلول A ، أما الثانية فقد عولجت بنفس الكمية السابقة بالمحلول B . وقد وجد أن المحلول A قد أهلك 420 بعوضة ، في حين أهلك المحلول B 390 بعوضة ، لنفتش عن تقدير الفرق بين قدرتي المحلولين على إبادة البعوض عند استخدامهما في نفس الشروط المحيطة .

نلاحظ أن تعرض كل بعوضة لهذا المحلول هو تكرار مستقل من مجتمع حداني . فإذا فرضنا أن احتمال النجاح ( أى هلاك بعوضة ) هو  $P_1$  بالنسبة للمحلول A و  $P_2$  بالنسبة للمحلول B فعندئذ يكون علينا إيجاد تقدير للفرق  $P_1 - P_2$  . باستخدام معطيات الحل نجد أن :

$$n_1 = 500 , X_1 = 420 , \hat{P} = \frac{X_1}{n_1} = 0.84$$

$$n_2 = 500 , X_2 = 390 , \hat{P} = \frac{X_2}{n_2} = 0.78$$

وكتقدير للفرق  $P_1 - P_2$  نأخذ التقدير التقطي  $\hat{P}_1 - \hat{P}_2 = 0.06$  لنفتش عن 0.95 مجال ثقة لهذا التقدير . من الواضح أن 0.95 مجال الثقة هو المجال :

$$\hat{P}_1 - \hat{P}_2 - Z_{\alpha/2} \cdot \sqrt{\frac{\hat{P}_1 \hat{Q}_1}{n_1} + \frac{\hat{P}_2 \hat{Q}_2}{n_2}} , \hat{P}_1 - \hat{P}_2 + Z_{\alpha/2} \cdot \sqrt{\frac{\hat{P}_1 \hat{Q}_1}{n_1} + \frac{\hat{P}_2 \hat{Q}_2}{n_2}}$$

وبالتعويض عن قيمة  $Z_{\alpha/2} = Z_{0.025} = 1.96$  من الجدول IV نجد أن 0.95 مجال الثقة لفرق التقديرين هو المجال :

$$(0.06 - 1.96 \sqrt{\frac{(0.84)(0.16)}{500} + \frac{(0.78)(0.22)}{500}} ,$$

$$(0.06 + 1.96 \sqrt{\frac{(0.84)(0.16)}{500} + \frac{(0.78)(0.22)}{500}} )$$

أما حدود خطأ هذا التقدير فهي :

$$(0.06 - 0.048, 0.06 + 0.048) = (0.012, 0.108)$$

وثقتنا يمثل هذا التقدير ناتجة عن معرفتنا بأنه إذا أعدنا نفس التجربة مرارا وتكراراً ، وحسبنا في كل مرة تقديراً مجالياً للفرق  $P_1 - P_2$  ، فإن 95% تقريباً من هذه المجالات ستحتوي القيمة الحقيقية للفرق  $P_1 - P_2$  .

### (٦,٧) تقدير التباين Estimating the variance

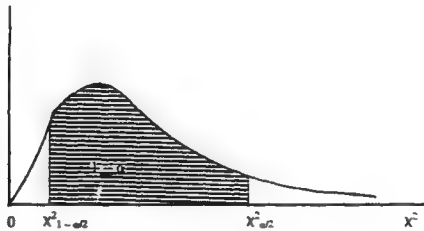
كتقدير نقطي لتباين مجتمع إحصائي  $\sigma^2$  نأخذ عادة تباين العينة  $S^2$  والذي يمثل كما أوضحنا تقديراً غير متحيز لتباين المجتمع المجهول  $\sigma^2$   $E(S^2 = \sigma^2)$  من المعلوم أن توزيع الإحصاء :

$$X^2 = \frac{(n-1) S^2}{\sigma^2}$$

هو التوزيع  $\chi^2_{(n-1)}$  (كاي - مربع) بـ  $(n-1)$  درجة من الحرية النظرية  $(5, 16)$  ، وذلك بفرض أن العينة التي حسبنا منها الإحصاء السابق قد أخذت من مجتمع طبيعي تباينه  $\sigma^2$  . لذلك يمكن أن نكتب :

$$P(X^2_{1-\alpha/2} < X^2 < X^2_{\alpha/2}) = 1 - \alpha$$

حيث تمثل  $X^2_{\alpha/2}$  قيمة كاي - مربع بـ  $(n-1)$  درجة حرية ، والتي تحدد على يمينها وتحت منحنى كثافة  $\chi^2$  مساحة قدرها  $\alpha/2$  انظر الشكل (٦,٤)



الشكل (٦,٤)

كما نلاحظ على الشكل (٦،٤) أن  $\chi^2_{1-\alpha/2}$  تمثل قيمة  $\chi^2$  بـ  $n-1$  درجة حرية ، والتي تحدد على يمينها مساحة قدرها  $1-\alpha/2$  .

وبتبديل قيمة الإحصاء  $S^2 = \frac{(n-1)}{\sigma^2} X^2$  فإننا نجد أن :

$$P \left[ \chi^2_{1-\alpha/2} < \frac{(n-1) S^2}{\sigma^2} < \chi^2_{\alpha/2} \right] = 1 - \alpha$$

وبتقسيم كافة عناصر المتباينة على  $(n-1)S^2$  ثم بأخذ مقلوب هذه المتباينة المزدوجة نجد أن :

$$P \left[ \frac{(n-1)S^2}{\chi^2_{\alpha/2}} < \sigma^2 < \frac{(n-1)S^2}{\chi^2_{1-\alpha/2}} \right] = 1 - \alpha$$

وأخيراً فإن  $100\%(1-\alpha)$  مجال ثقة للتباين المجهول  $\sigma^2$  في مجتمع إحصائي هو المجال :

$$\left( \frac{n-1}{\chi^2_{\alpha/2}} S^2 , \frac{n-1}{\chi^2_{1-\alpha/2}} S^2 \right)$$

مثال (٦،١٤)

سحبنا عينة عشوائية حجمها  $n = 20$  من مجتمع طبيعي . فأعطت وسط عينة  $\bar{x} = 32.8$  وتباين عينة  $S = 4.51$  . لنحسب  $0.95$  مجال ثقة للتباين  $\sigma^2$  .

الحل

نلاحظ أن  $1 - \alpha = 0.95$  . أي أن  $\frac{\alpha}{2} = 0.025$  . كما نلاحظ أنه من أجل  $n-1 = 19$   $\chi^2_{0.025} = 32.852$  وأن  $\chi^2_{0.975} = 8.907$  . كذلك فإن  $S = 4.51$  . وهكذا نجد أن  $0.95$  مجال ثقة للتباين  $\sigma^2$  هو المجال :

$$\left( \frac{n-1}{\chi^2_{\alpha/2}} S^2 , \frac{n-1}{\chi^2_{1-\alpha/2}} S^2 \right) \left( \frac{19}{32.852} 4.51 , \frac{19}{8.907} 4.51 \right)$$

أي المجال :

$$(2.608 , 9.620)$$

### Estimating the Ratio of two variances (٦,٨) تقدير نسبة تباينين

كتقدير نقطي لنسبة تباينين مجتمعين إحصائيين  $\frac{\sigma_1^2}{\sigma_2^2}$  نأخذ عادة النسبة  $\frac{S_1^2}{S_2^2}$  أى نسبة تباينين العيتين المسحوبتين من هذين المجتمعين . لذلك فإن الإحصاء  $\frac{S_1^2}{S_2^2}$  يشكل تقديراً لـ  $\frac{\sigma_1^2}{\sigma_2^2}$  . إذا كان المجتمعان المدروسان طبيعيين فعندئذ يمكن إنشاء مجال لتقدير  $\frac{\sigma_1^2}{\sigma_2^2}$  وذلك باستخدام الإحصاء .

$$F = \frac{\sigma_2^2 S_1^2}{\sigma_1^2 S_2^2}$$

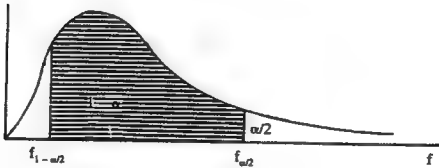
وبحسب النظرية (٥,٢٠) فإن المتغير العشوائى F له توزيع F بعدد من درجات الحرية  $\nu_2 = n_2 - 1$  ،  $\nu_1 = n_1 - 1$  لذلك واعتماداً على الشكل (٦,٥) يمكننا أن نكتب :

$$P [ f_{1-\alpha/2} (\nu_1, \nu_2) < F < f_{\alpha/2} (\nu_1, \nu_2) ] = 1 - \alpha$$

حيث تمثل  $f_{\alpha/2} (\nu_1, \nu_2)$  قيمة F بـ  $\nu_1, \nu_2$  درجات حرية ، والتي تحدد على يمينها مساحة قدرها  $\frac{\alpha}{2}$  ، انظر الشكل (٦,٥) .

وبتعبير قيمة F فى العلاقة السابقة نجد أن :

$$P \left[ f_{1-\alpha/2} (\nu_1, \nu_2) < \frac{\sigma_2^2 S_1^2}{\sigma_1^2 S_2^2} < f_{\alpha/2} (\nu_1, \nu_2) \right] = 1 - \alpha$$



الشكل (٦,٥)



أى أن :

$$P \left[ \frac{S_1^2}{S_2^2} \frac{1}{f_{\alpha/2}(v_1, v_2)} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{S_1^2}{S_2^2} \frac{1}{f_{1-\alpha/2}(v_1, v_2)} \right] = 1 - \alpha$$

وحيث إن  $f_{1-\alpha/2}(v_1, v_2) = \frac{1}{f_{\alpha/2}(v_2, v_1)}$  لذلك فإننا نجد :

$$P \left[ \frac{S_1^2}{S_2^2} \frac{1}{f_{\alpha/2}(v_1, v_2)} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{S_1^2}{S_2^2} f_{\alpha/2}(v_2, v_1) \right] = 1 - \alpha$$

وهكذا نستنتج بأنه  $100\%(1 - \alpha)$  مجال ثقة للنسبة  $\frac{\sigma_1^2}{\sigma_2^2}$  هو المجال :

$$\frac{S_1^2}{S_2^2} \frac{1}{f_{\alpha/2}(v_1, v_2)} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{S_1^2}{S_2^2} f_{\alpha/2}(v_2, v_1)$$

وذلك بفرض أننا سحبنا عيّتين مستقلتين ذواتا حجمين  $n_1, n_2$  من مجتمعين طبيعيين لهما تباينان  $\sigma_1^2$  و  $\sigma_2^2$  .

## مثال (٦,١٥)

أقامت جامعة الملك عبد العزيز بمكة مسابقة بين طلاب كليتي العلوم والهندسة وتقدم 25 طالبا من كلية العلوم إلى المسابقة ، ومن كلية الهندسة 16 طالبا وفي اختبار الرياضيات قدم طلاب كلية العلوم هذا الاختبار بوسط قدره 82 وانحراف معياري 8 . بينما قدم طلاب كلية الهندسة الاختبار بوسط 78 وانحراف معياري قدره 7 . لننشئ 98% مجال ثقة للنسبة  $\frac{\sigma_1^2}{\sigma_2^2}$  ثم  $\frac{\sigma_1}{\sigma_2}$  ، وذلك بفرض أن  $\sigma_1^2$  و  $\sigma_2^2$  يمثلان تبايني مجتمعى درجات كل من طلاب كليتي العلوم والهندسة على الترتيب والمتقدمين إلى مثل هذا الاختبار .

نلاحظ أن :

$$n_1 = 25, n_2 = 16, S_1 = 8, S_2 = 7$$

ومن أجل 98% مجال ثقة نلاحظ أيضاً أن  $\alpha = 0.02$  . وباستخدام الجدول VII نجد أن :

$$f_{0.01}(24, 15) = 3.29 \quad , \quad f_{0.01}(15, 24) = 2.89$$

وبالتبديل في العلاقة :

$$\frac{S_1^2}{S_2^2} \frac{1}{f_{\alpha/2}(v_1, v_2)} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{S_1^2}{S_2^2} f_{\alpha/2}(v_2, v_1)$$

فإننا نجد أن :

$$\frac{(64)}{(49)} \frac{1}{(3.29)} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{(64)}{(49)} (2.89)$$

وأخيراً فإن 0.98 مجال ثقة لنسبة التباين هو المجال :

$$0.397 < \frac{\sigma_1^2}{\sigma_2^2} < 3.775$$

وبأخذ جذرى أطراف المتباينة السابقة نحصل على 0.98 مجال ثقة لـ  $\frac{\sigma_1}{\sigma_2}$  هو المجال :

$$0.630 < \frac{\sigma_1}{\sigma_2} < 1.943$$

## تمارين محلولة

## تمرين (٩)

بفرض أن وسط العينة والانحراف المعياري المحسوبين من خلال عينة حجمها  $n = 36$  مسحوبة من مجتمع من الرجال البالغين هما على الترتيب 2.6 , 0.3 . أوجد 95% ثم 99% مجالات ثقة للوسط  $\mu$  المجهول في هذا المجتمع .

## الحل

نلاحظ أن  $\bar{x} = 2.6$  هو تقدير نقطي للوسط المجهول  $\mu$  في هذا المجتمع ، وبما أن حجم العينة المسحوبة  $n = 36$  ، لذلك يمكن أن نقرب التباين  $\sigma^2$  إلى  $S^2$  المحسوبة من خلال العينة واعتبار أن  $\sigma = S = 0.3$  .

أما بالنسبة للنقطة  $Z_{\alpha/2}$  فهي النقطة من محور السينات التي تحصر لنا على الترتيب مساحة قدرها 0.025 على يمينها أو 0.975 على يسارها من أجل 95% ثقة ، و 0.005 على يمينها أو 0.995 على يسارها من أجل 99% . وهكذا نجد أن 95% مجال الثقة للوسط  $\mu$  هو المجال :

$$2.6 - (1.96) \left( \frac{0.3}{\sqrt{36}} \right) < \mu < 2.6 + (1.96) \left( \frac{0.3}{\sqrt{36}} \right)$$

أى أن :

$$2.5 < \mu < 2.71$$

كما نلاحظ أن 99% مجال ثقة لنفس الوسط هو المجال :

$$2.6 - (2.575) \left( \frac{0.3}{\sqrt{36}} \right) < \mu < 2.6 + (2.575) \left( \frac{0.3}{\sqrt{36}} \right)$$

أى أن :

$$2.47 < \mu < 2.73$$

وهكذا نستنتج أن مجال الثقة الأطول يعطى تقديراً بدرجة أعلى من الدقة .

تمرين (٢)

$$\cdot \hat{S}^2 = \frac{n}{\sum_{i=1}^n (X_i - \bar{X})^2} / n$$

$$\cdot E(\hat{S}^2) = \frac{n-1}{n} \sigma^2$$

وأن  $S^2$  هو تقدير متحيز لـ  $\sigma^2$  .

الحل

من الواضح أن :

$$\frac{n \hat{S}^2}{n-1} = \sum_{i=1}^n (X_i - \bar{X})^2 / (n-1)$$

$$\frac{n \hat{S}^2}{n-1} = S^2$$

ثم أن :

$$E \left( \frac{n \hat{S}^2}{n-1} \right) = E(S^2)$$

وبما أن  $S^2$  هو تقدير غير متحيز لـ  $\sigma^2$  مثال (٦،٢) لذلك فإن :

$$E \left( \frac{n \hat{S}^2}{n-1} \right) = \sigma^2$$

وهكذا نجد أن :

$$E \hat{S}^2 = \left( \frac{n-1}{n} \right) \sigma^2$$

وحيث أن :  $E(\hat{S}_2) = \sigma^2$   
لذلك لا يمثل  $S^2$  تقديراً غير متحيز لـ  $\sigma^2$  فهو يمثل تقديراً متحيزاً لهذا التباين .

### تمرين (٣)

المطلوب تقدير وسط الإنتاج اليومي في منشأة للصناعات الكيماوية ، وذلك بفرض أن وسط عينة من الإنتاج اليومي وانحرافها المعياري لفترة  $n = 50$  يوماً هما على التتالي :

$$\bar{X} = 871$$

$$S = 21$$

### الحل

من المعلوم أن التقدير الأفضل هو  $\bar{X} = 871$  طن يومياً ، وحدود الخطأ في هذا التقدير هي :

$$\pm 2\sigma_{\bar{x}} = \pm 2 \frac{\sigma}{\sqrt{n}} = 2 \frac{\sigma}{\sqrt{50}}$$

ومع أن  $\sigma$  مجهول ، إلا أنه يمكن اعتبار  $S$  كقيمة تقريبية له أى اعتبار  $S$  تقديراً لـ  $\sigma$  .  
وهكذا يكون حدود الخطأ في هذا التقدير بصورة تقريبية مساوياً لـ :

$$\frac{2S}{\sqrt{50}} = \frac{2(21)}{\sqrt{50}} = 5.94$$

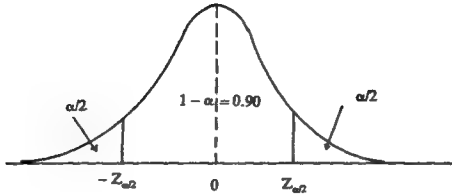
وسنشعر إلى حد ما بالثقة من أن التقدير 871 هو في حدود 5.94 طن من القيمة الحقيقية لوسط الإنتاج .

### تمرين (٤)

احسب 90% مجال ثقة لوسط الإنتاج اليومي في التمرين (٦،٣) .

الحل

نلاحظ أن :



إن  $Z_{\alpha/2}$  الموافقة لأمثال ثقة 90% هي القيمة  $Z_{\alpha/2} = 1.645$  ، ومنه فمجال الثقة المطلوب هو المجال :

$$\bar{x} \pm 1.645 \frac{\sigma}{\sqrt{n}}$$

فإذا وضعنا S بدلا عن  $\sigma$  فإننا نحصل بصورة تقريبية على المجال :

$$871 \pm (1.645) \frac{21}{\sqrt{50}}$$

أو :

$$871 \pm 4.89$$

وهكذا نقدر أن وسط الإنتاج اليومي  $\mu$  يقع ضمن المجال من 866.11 إلى 875.89 طنا .  
إن أمثال الثقة السابقة تعنى أن 90% من مجالات الثقة الناتجة عن سحب عينات بصورة متكررة ستحوى  $\mu$  .

تمرين (٥)

احسب 95% مجال ثقة لوسط سعة أو عينة تحتوي على حمض الكبريت إذا علمت أن سبعة منها متشابهة وتتسع على التتالي :

9.8 , 10.2 , 10.4 , 9.8 , 10 , 10.2 , 9.6

الحل

نلاحظ أن وسط العينة والانحراف المعياري لهذه العينة من الأوعية هما على التالي :

$$\bar{x} = 10.0 , \quad s = 0.283$$

باستخدام الجدول  $v$  نجد أن  $t_{0.025} = 2.447$  من أجل  $n - 1 = 7 - 1 = 6$  درجة من الحرية . لذلك فإن 95% مجال ثقة للوسط  $\mu$  ( الممثل لوسط سعة الأوعية المستخدمة ) هو المجال :

$$10.0 - (2.447) \left( \frac{0.283}{\sqrt{7}} \right) < \mu < 10.0 + (2.447) \left( \frac{0.283}{\sqrt{7}} \right)$$

$$9.74 < \mu < 10.26$$

تمرين (٦)

في الاختبار النهائي لمادة الاحتمالات تقدمت مجموعتان تحويان 75 طالبا من الشعبة الأولى و 50 طالبا من الشعبة الثانية . وقد قدم طلاب المجموعة الأولى هذا الامتحان بوسط 82 درجة وانحراف معياري قدره 8 درجة ، أما طلاب المجموعة الثانية فقد قدموا امتحانهم هذا بوسط قدره 76 درجة وانحراف معياري 6 . والمطلوب إيجاد 96% مجال ثقة للفرق  $\mu_1 - \mu_2$  حيث يمثل  $\mu_1$  وسط الدرجات التي أحرزها كل طلاب الشعبة الأولى ، أما  $\mu_2$  فتمثل وسط الدرجات التي أحرزها كل طلاب الشعبة الثانية . لاحظ الشكل التالي .



## الحل

من شروط المسألة نجد أن :

$$\bar{x}_1 = 82 \quad s_1 = 8$$

$$\bar{x}_2 = 76 \quad s_2 = 6$$

نلاحظ أن التقدير النقطي لفرق الوسطين  $\mu_1 - \mu_2$  هو :  $\bar{x}_1 - \bar{x}_2 = 82 - 76 = 6$  .  
كما نلاحظ أن حجم العينة الأولى  $n_1 = 75$  ، وحجم العينة الثانية  $n_2 = 50$  . لذلك يمكن استبدال  $\sigma_1$  و  $\sigma_2$  على الترتيب بـ  $s_1$  و  $s_2$  وحيث إن  $1 - \alpha = 0.96$  لذلك  $\alpha = 0.04$  . ومن الجدول IV نجد أن  $Z_{\alpha/2} = Z_{0.02} = 2.054$  . وهكذا نجد أن 96% مجال ثقة لفرق الوسطين هو المجال :

$$\left( (\bar{x}_1 - \bar{x}_2) - Z_{\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} , (\bar{x}_1 - \bar{x}_2) + Z_{\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \right)$$

أى :

$$\left( 6 - 2.054 \sqrt{\frac{64}{75} + \frac{36}{50}} , 6 + 2.054 \sqrt{\frac{64}{75} + \frac{36}{50}} \right)$$

$$3.42 < \mu_1 - \mu_2 < 8.58$$

## تمرين (٧)

عينة عشوائية مؤلفة من قياسات قطر جسم كروي كانت قد سجلت على النحو

التالى :

$$3.37 , 6.36 , 6.37 , 6.35 , 6.34 \text{ mm}$$

أوجد تقديراً غير متحيز للوسط المجهول ثم للتباين المجهول أيضا .

## الحل

نلاحظ أن وسط العينة يشكل تقديراً غير متحيز للوسط المجهول ، ولذلك فإن :



$$\bar{x} = \frac{\sum_{i=1}^n x_i}{5} = \frac{6.37 + 6.36 + 6.37 + 6.35 + 6.34}{5} = 6.35 \text{ mm}$$

ونعلم أن تباين العينة  $S^2$  يشكل تقديراً غير متحيز للتباين  $\sigma^2$  كما أن :

$$S^2 \approx \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{5 - 1}$$

$$\frac{(6.37 - 6.35)^2 + (6.36 - 6.35)^2 + (6.37 - 6.35)^2 + (6.34 - 6.35)^2 + (6.35 - 6.35)^2}{4} \\ = S^2$$

$$S^2 = 0.00025 (\text{mm})^2$$

### تمرين (٨)

لقياس زمن رد الفعل ، قدر أحد العلماء الانحراف المعياري بـ 0.05 ثانية . ما هو حجم العينة من القياسات بحيث تكون بـ 95% . واثقين من أن خطأ التقدير لن يتجاوز 0.01 ثانية ؟

### الحل

$$1 - \alpha = 0.95$$

من الفرض لدينا

$$Z_{\alpha/2} = Z_{0.025} = 1.96$$

ولذلك فإن :

$$\sigma = S = 0.05 \text{ ثانية}$$

كما أن :

ونعلم أن الخطأ  $e$  يعطى بالعلاقة :

$$e \leq \frac{\sigma}{\sqrt{n}} Z_{\alpha/2}$$

ومن هنا نجد أن :

$$n \leq \left( \frac{\sigma}{e} Z_{\alpha/2} \right)^2$$

$$n \leq \left( 1.96 \frac{0.05}{0.01} \right)^2$$

$$n \leq 96.04$$

أى أن :

$$n = 96$$

### تمرين (٩)

تم استطلاع الرأى العام فى جامعة الملك عبد العزيز حول انتخاب المرشح محمد بن المبارك لشغل منصب أمين عام العلاقات العامة . فاختيرت عينة من المسئولين ممن يحق لهم الانتخاب حجمها 100 شخص وتم استجوابهم حول ترشيحهم للسيد المذكور فدلّت النتائج على أن 57% منهم يؤيدون ترشيح المرشح المذكور . أوجد 99% مجال ثقة للنسبة بين جميع الناخبين المؤيدين للسيد محمد بن المبارك .

### الحل

نفرض أن P يمثل نسبة المؤيدين للمرشح . ومن الفرض لدينا :

$$1 - \alpha = 0.99$$

$$\alpha = 0.01 \rightarrow Z_{\alpha/2} = Z_{0.005}$$

لذلك فإن :

نلاحظ من الجدول IV أن النقطة  $Z_{0.005}$  التى تحصر على يسارها تحت منحنى الكثافة الطبيعية المعيارى مساحة قدرها 0.995 هى النقطة :

$$Z_{\alpha/2} = 2.58$$

$$\hat{p} = 0.55$$

ثم إن :

$$\hat{q} \geq 1 - \hat{p} = 0.45$$

$$n = 100$$

ونعلم أن 100% (1 -  $\alpha$ ) مجال الثقة للنسبة P هو المجال :

$$\left( \hat{p} - Z_{\alpha/2} \sqrt{\frac{\hat{p}\hat{q}}{n}}, \hat{p} + Z_{\alpha/2} \sqrt{\frac{\hat{p}\hat{q}}{n}} \right)$$

وبالتعويض نجد أنه :

$$(0.55 - 2.58) \sqrt{\frac{(0.55)(0.45)}{100}} , 0.55 + 2.58 \sqrt{\frac{(0.55)(0.45)}{100}} )$$

وأخيراً فإن مجال الثقة هو :

$$(0.4216 , 0.6783)$$

تمرين (١٠)

يُنتج مصنعان A و B نوعين متشابهين من اللبمبات الكهربائية . أخذت عينة من إنتاج المصنع A حجمها  $n_1 = 150$  لمبة وجربت ، فكان وسط عمر كل لمبة من هذه اللبمبات 1400 ساعة عمل وانحرافها المعياري 120 ساعة ، كما أخذت عينة أخرى إنتاج المصنع B حجمها  $n_2 = 200$  لمبة وجربت وسجلت نتائج التجربة . فإذا علمت أن وسط عمر كل لمبة كان 1200 ساعة وانحرافها المعياري 80 ساعة ، فأوجد 95% مجال ثقة للفرق بين وسط العمر لللبمبات المصنعين A و B .

الحل

$$1 - \alpha = 0.95 \quad \text{نعلم أن :}$$

$$Z_{\alpha/2} = 1.96 \quad \text{وهذا يعنى :}$$

$$\bar{x}_1 = 1400 , \sigma_1^2 = (120)^2 \quad \text{لدينا أيضاً :}$$

$$\bar{x}_2 = 1200 , \sigma_2^2 = (80)^2$$

فإذا فرضنا أن وسط عمر أية لمبة من إنتاج المصنع A هو  $\mu_1$  وأن  $\mu_2$  هو وسط عمر أى لمبة من إنتاج المصنع B . فمن المعلوم أن 95% مجال ثقة لفرق الوسطين  $\mu_1 - \mu_2$  هو المجال :

$$\left( \bar{x}_1 - \bar{x}_2 - \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \cdot Z_{\alpha/2} , \bar{x}_1 - \bar{x}_2 + \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \cdot Z_{\alpha/2} \right)$$

وبالتعويض نجد أن :

$$1200 - 1.96 \sqrt{\frac{(120)^2}{150} + \frac{(80)^2}{200}} , 1400 - 1200 + 1.96 \sqrt{\frac{(120)^2}{150} + \frac{(80)^2}{200}} \\ (1400 - \\ \text{أى أن :}$$

تقرين (١١)

عشرة رزم من حبوب الأعشاب أوزانها بالديسغرام :

$$46.9 , 45.2 , 46.4 , 46.0 , 46.1 , 45.8 , 47.0 , 46.1 , 45.9 , 45.8$$

أوجد 0.95 مجال ثقة لتباين كل رزم الأعشاب السابقة .

الحل

$$S^2 = \frac{n \sum_{i=1}^n X_i^2 - (\sum_{i=1}^n X_i)^2}{n(n-1)} = \frac{10 (21273.12) - (461.2)^2}{(10)(9)} = 0.286$$

وللحصول على 0.95 مجال ثقة نلاحظ أن  $\alpha = 0.05$  . ومن الجدول IV نجد أن  $\chi^2_{0.025} = 19.023$  وأن  $\chi^2_{0.975} = 2.70$  وذلك من أجل  $v = 9$  ، وبالتعويض في العلاقة :

$$\frac{(n-1)S^2}{\chi^2_{\alpha/2}} < \sigma^2 < \frac{(n-1)S^2}{\chi^2_{1-\alpha/2}}$$

نجد أن :

$$\frac{(9)(0.286)}{19.023} < \sigma^2 < \frac{(9)(0.286)}{2.70}$$

$$0.135 < \sigma^2 < 0.953$$

## تمارين عامة

(١) ينتج مصنعا للمصابيح الكهربائية نوعا من النيون ، وبفرض أن  $\mu$  يمثل وسط عمر أى نيون من إنتاج المصنع ، وبقصد البحث عن تقدير لوسط العمر  $\mu$  اخترنا بصورة عشوائية عينة من مئتي نيون فأعطت وسط عمر قدره 1500 ساعة وانحراف معيارى قدره 143 ساعة ، والمطلوب البحث عن تقدير غير متحيز لـ  $\mu$  ، ووضع حدود لخطأ هذا التقدير .

(٢) اخترنا عينة عشوائية مؤلفة من 300 ترانزيستور وجربناها ، فوجدنا أنها تحتوى على 30 ترانزيستورا عاطلا . أنشئ 0.90 مجال ثقة لنسبة الترانزيستورات العاطلة .

(٣) إذا علمت أن وسط عينة مؤلفة من 50 قياسا هو 36.5 ، وأن انحرافها المعيارى هو 2.1 . فالمطلوب إنشاء 0.98 مجال ثقة لوسط المجتمع الذى أخذت منه هذه العينة ، ضع حدودا لخطأ هذا التقدير .

(٤) ترغب إدارة مستشفى جامع الملك عبد العزيز بمجدة في تقدير عدد الأيام التى يحتاجها علاج مرضى الجرب ، والذين يتراوح سنهم بين 15 و 50 عاما . ولهذا الغرض اختيرت عينة عشوائية من المرضى مؤلفة من 500 مريضاً . فوجد أن وسط عدد الأيام في هذه العينة هو 5.3 يوما ، وأن انحرافها المعيارى هو 1.25 يوما . والمطلوب تقدير وسط الإقامة في المستشفى المذكور لجميع مرضى الجرب الذين اختيرت منهم هذه العينة .

(٥) استخدم عمال الحفر للتنقيب عن البترول نوعين من الحفارات من إنتاج مصنعين مختلفين A و B ، وقد وجدوا أن وسط اختراق الحفارة الأولى ( من إنتاج المصنع A ) هو 10.8 ، وأن انحرافها المعيارى هو 1.2 بوصة . وذلك عندما قاموا بحفر 50 حفرة . كما وجدوا أن وسط اختراق الحفارة من النوع الثانى هو 9.1 وانحرافها المعيارى 1.6 بوصة وذلك عند حفر 80 حفرة صخرية أخرى . والمطلوب تقدير الفرق في معدل الاختراق الوسطى لهاتين الحفارتين إذا علمت أن بنية التربة ومدة الحفر واحدة بالنسبة للحفارتين .

(٦) تم سحب عيتين مستقلتين حجم كل منهما  $n = 64$  من مجتمعين طبيعيين لهما نفس الوسط 6.4 ونفس الانحراف المعياري 7.2 . ما هو احتمال أن يتجاوز الفرق بين وسطي العينين العدد 90.6 بالقيمة المطلقة ؟

(٧) اختبرت عيتان عشوائيتان حجمهما  $n_1 = 9$  ,  $n_2 = 16$  من مجتمعين طبيعيين مستقلين بالوسطين العينين  $\bar{x}_1 = 64$  ,  $\bar{x}_2 = 59$  والانحرافين المعياريين  $S_1 = 6$  ,  $S_2 = 5$  . أنشئ 95% مجال ثقة للفرق  $\mu_1 - \mu_2$  وذلك بفرض أن  $\sigma_1 = \sigma_2$  .

(٨) في مدينة معينة اختبرت عينة من 200 شخص بصورة عشوائية . ووجد أن 114 منهم يؤيدون ترشيح المرشح X . أنشئ 96% مجال ثقة لنسبة المؤيدين لهذا المرشح في المدينة المذكورة .

(٩) اختبر 500 شخص من المدخنين في مدينة جدة بالملكة العربية السعودية وتم سؤالهم عن رأيهم حول نوع معين من السجائر فوجد أن 86 شخصاً منهم يفضلون هذا النوع . أنشئ 90% مجال ثقة لنسبة المدخنين الذين يفضلون هذا النوع من السجائر .

(١٠) سحبنا عينة عشوائية حجمها  $n = 100$  عنصر من جملة بضاعة مصنعة جاهزة للتسويق وفحصناها فوجدنا أن هذه العينة تحتوي على ثمانية عناصر معيبة ، أنشئ 98% مجال ثقة لنسبة العناصر المعيبة .

(١١) تحاول شركة لتأجير السيارات شراء إطارات لسيارتها ( التي تعمل بين مكة المكرمة — جدة ) من أحد النوعين A و B من الإطارات ، ولتقدير الفرق بين النوعين اختارت هذه الشركة اثني عشر إطاراً من كل نوع ووضعتهما تحت الخدمة حتى بليت جميعها . وقد أعطت التجربة السابقة المعلومات التالية حول هذين النوعين ، فبالنسبة للنوع A أوضحت التجربة أن ( كيلومتراً  $S_1 = 5000$  ) وأن ( كيلومتراً  $\bar{x}_1 = 36300$  ) أما بالنسبة للنوع B فأوضحت أن ( كيلومتراً  $S_2 = 6100$  ) ، وأن ( كيلومتراً  $\bar{x}_2 = 38100$  ) والمطلوب إنشاء 95% مجال ثقة للفرق  $\mu_1 - \mu_2$  ، وذلك بفرض أن لكل من مجتمعى الإطارات السابقة توزيعاً قريباً من التوزيع الطبيعي .

(١٢) أنشئ 90% مجال ثقة لـ  $\sigma_1^2 / \sigma_2^2$  في المثال السابق .

## الفصل السَّابع

### إختبارات الفرضيات

- الفرضية الاحصائية
- الأخطاء من النوع الأول I ومن النوع الثاني II
- الإختبارات وحيدة وثنائية الذيل
- إختبار وسط وتباين مجتمع احصائي
- إختبار حجم العينة لإختبار الوسط
- الإختبارات المتعلقة بالنسب
- إختبار الفرق بين نسبتين
- تقاوين معلولة.





## (٧,١) الفرضية الإحصائية

كثيراً ما نلجأ إلى اتخاذ قرار يخص مجتمعاً مدروساً وذلك بناء على معلومات تقدمها العينة الإحصائية ، مثل هذا القرار يدعى عادة بالقرار الإحصائي . وبغية الحصول على قرار مفيد لابد من وضع فروض أو تخمينات تتعلق أيضاً بهذا المجتمع الذي ندرسه ، مثل هذه الفروض قد تكون صحيحة أو غير صحيحة وتدعى بالفروض الإحصائية .

ويتألف أى اختبار إحصائي من أربعة عناصر :

- ١ - فرض العدم
- ٢ - إحصاء الاختبار
- ٣ - منطقة الرفض
- ٤ - الفرضية البديلة

إن تحديد هذه العناصر الأربعة اختباراً معيناً ، وتغيير عنصر أو أكثر يؤدي إلى اختبارات جديدة .

## (٧,٢) الأخطاء من النوع الأول I ومن النوع الثاني II

Errors: kind I and II

نرمز عادة لفرض العدم بالرمز  $H_0$  ، وهذا الفرض يحدد قيمة افتراضية لوسيط أو أكثر من وسطاء المجتمع . فقد نرغب مثلاً في اختبار الفرض بأن نوعاً معيناً من لقاح الجدرى الملاء فعالاً خلال عامين بنسبة 25% .

إن قرارنا برفض أو قبول فرض العدم  $H_0$  مبني على المعلومات التي نحويها عينة مسحوبة من المجتمع الإحصائي المدروس . وتستخدم قيم العينة لحساب عدد واحد يأخذ

دور صانع القرار ، ويدعى بإحصاء الاختبار . ونقسم مجموعة كل القيم التي يمكن أن يأخذها إحصاء الاختبار إلى مجموعتين أو منطقتين تدعى إحداهما بمنطقة الرفض ، كما تدعى الثانية بمنطقة القبول . فإذا وقعت القيمة التي يأخذها إحصاء الاختبار في منطقة الرفض ، فإننا نرفض الفرضية الابتدائية . ونقبلها إذا وقعت في منطقة القبول .

وتخضع طريقة اتخاذ القرار هذه لنوعين من الخطأ . إذ يمكن أن نرفض فرض العدم ، بينما هو في الواقع صحيح ، أو يمكن أن نقبل  $H_0$  وهو في الحقيقة غير صحيح ، والصحيح هو فرض آخر بديل . يدعى هذان النوعان من الأخطاء بالخطأ من النوع الأول I ، والخطأ من النوع الثاني II على الترتيب . أى أن رفض فرض صحيح ( أو الرفض الخاطئ ) هو الخطأ من النوع الأول ، وقبول فرض غير صحيح ( أى القبول الخاطئ ) هو الخطأ من النوع الثاني . ويوضح الجدول (٧،١) الحالتين الممكنتين لفرض العدم ، ونوعى القرار الممكنين ، بالإضافة إلى نوع الخطأ الذي يمكن ارتكابه .

القرار	فرض العدم	
	صحيح	مخطئ
الرفض	الخطأ من النوع الأول I	قرار صحيح
القبول	قرار صحيح	الخطأ من النوع الثاني II

الجدول (٧،١)

ونقتبس جودة الاختبار الإحصائي لفرض باحتمالي الخطأ من النوع الأول I ، والخطأ من النوع الثاني II ، ونرمز لهما بالرمزين  $\alpha$  ،  $\beta$  على التوالي . ويمكن التعبير عن الاحتمال  $\alpha$  بأنه احتمال أن يقع إحصاء الاختبار في منطقة الرفض علماً بأن  $H_0$  صحيح ، تدعى  $\alpha$  أيضاً بمستوى معنوية الاختبار . ومن الواضح أن زيادة حجم منطقة الرفض سيزيد من قيمة  $\alpha$  . وفي نفس الوقت سيؤدي إلى تناقص  $\beta$  . وذلك من أجل قيمة ثابتة لـ  $n$  ، أما تخفيض حجم منطقة الرفض فسيؤدي إلى تناقص  $\alpha$  وذلك من أجل قيمة ثابتة لـ  $n$  ، أما تخفيض حجم منطقة الرفض فسيؤدي إلى تناقص  $\alpha$  وزيادة  $\beta$  ، وذلك عند زيادة حجم العينة  $n$  . فمن الطبيعي أن تقدم لنا العينة الأكبر حجماً ، قدرنا أكبر من المعلومات نتخذ على ضوءها قرارنا ، وهذا يؤدي إلى تناقص كل من  $\alpha$  و  $\beta$  .

لنفرض على سبيل المثال أن لقاحاً معيناً ضد جدري الماء فعال بنسبة 20% بعد عامين من استخدامه ، لنفرض أن شركة للأدوية اكتشفت لقاحاً جديداً ضد نفس المرض ، وتريد معرفة فعالية هذا الدواء لفترة زمنية أكبر . لذلك اختارت هذه الشركة بصورة عشوائية 15 شخصاً وتم تلقيحهم بهذا اللقاح الجديد . وتعتبر الشركة أن اللقاح الجديد هو أكثر فعالية إذا صدف أن سبعة أو أكثر من هؤلاء الأشخاص قد مر على تلقيحهم عامين أو أكثر دون التقاطهم لفيروس هذا المرض .

نلاحظ أن العدد 7 كفي إلى حد ما ، ولكن يبدو معقولاً لأنه يمثل عدداً أكبر من ثلاثة وهو عدد الأشخاص الذين يتوقع أن يحصلوا على مقاومة ضد المرض فيما إذا تم تلقيح 15 شخصاً باللقاح القديم . لنختبر الآن الفرض العدم في أن المتغير الحداني ( وهو احتمال النجاح في اختبار ما ) هو  $P = \frac{1}{5}$  ضد الفرض البديل في أن  $P > \frac{1}{5}$  وهذا ما يكتب عادة على النحو التالي :

$$H_0 : P = \frac{1}{5}$$

$$H_0 : P > \frac{1}{5}$$

ويمثل الإحصاء  $X$  الذى بنى عليه قرارنا عدد الأشخاص في عينتنا ، والذين حصلوا باستخدام اللقاح الجديد على مقاومة لمدة عامين أو أكثر . لنجزى مجموعة القيم الممكنة من صفر إلى 15 إلى مجموعتين : تمثل الأولى مجموعة الأعداد الأقل من 7 ، أما الثانية فتمثل مجموعة الأعداد الأكبر من سبعة أو تساويها ، هذا وتؤلف مجموعة النقاط الأكبر من 6.5 منطقة تدعى بالمنطقة الحرجة ، كما تمثل مجموعة النقاط الأقل من 6.5 منطقة القبول 9 ، والعدد 6.5 الذى يفصل هاتين المنطقتين فيدعى بالقيمة الحرجة .

فإذا وقع الإحصاء  $X$  في المنطقة الحرجة فإننا نرفض  $H_0$  لصالح  $H_1$  ، أما إذا وقع  $X$  في منطقة القبول فإننا نقبل  $H_0$  .

نلاحظ أن احتمال ارتكاب خطأ من النوع الأول  $\alpha$  ، وهو ما سميناه بمستوى معنوية الاختبار يحسب على النحو التالي :

$$\alpha = P ( \text{الخطأ من النوع الأول} )$$

$$= P \left( X \geq 7 \mid P = \frac{1}{5} \right)$$

$$\begin{aligned}
 &= \sum_{x=7}^{15} b\left(x; 15, \frac{1}{5}\right) \\
 &= 1 - \sum_{x=0}^6 b\left(x; 15, \frac{1}{5}\right) \\
 &= 1 - 0.9819 \\
 &= 0.0181
 \end{aligned}$$

وهكذا فإننا نختبر فرض العلم  $P = \frac{1}{5}$  بمستوى معنوية قدره  $(\alpha = 0.0181)$

**ملاحظة :**

يسمى مستوى المعنوية في بعض الأحيان بحجم المنطقة الحرجة . والنتيجة السابقة تعنى المنطقة الحرجة ذات الحجم  $\alpha = 0.0181$  صغيرة جداً ، وهذا يعنى أنه من غير المحتمل في أننا سنرتكب خطأً من النوع الأول I . ومن غير الممكن حساب احتمال ارتكاب خطأً من النوع الثاني II قبل أن نحدد الفرض البديل تماماً . لنفرض مثلاً أن الفرض البديل هو  $P = \frac{1}{4}$  ، ففى هذه الحالة فإننا سنكون متمكنين من حساب .

$\beta = p$  ( احتمال ارتكاب خطأً من النوع الثانى )

$$\beta = P\left(X < 7 \mid P = \frac{1}{4}\right) \quad \text{ومنه :}$$

$$= \sum_{x=0}^6 b\left(x; 15, \frac{1}{4}\right)$$

$$= 0.9434 \quad \text{ومنه :}$$

ويشير هذا الاحتمال الكبير جداً إلى حد ما إذا ما قورن به  $\alpha = 0.0181$  إلى ضعف عملية الاختبار ، وهو يعنى أننا سنرفض اللقاح الجديد الذى يفوق اللقاح المستخدم حالياً في فعاليته . نرغب في كثير من الأحيان في استخدام إجراءات اختبار يكون معها كلا من  $\alpha$  و  $\beta$  صغيرين . وهذا ممكن بالطبع وهو بيد من يجرى عملية الاختبار السابقة .

لنفرض مثلاً أن  $P = 0.6 : H_1$  في هذه الحالة نلاحظ أن :

( الخطأ من النوع الثاني )  $\beta = P$

$$= P(X < 7 | P = 0.6)$$

$$= \sum_{x=0}^6 b(x; 15, 0.6)$$

$$= 0.0951$$

وبهذا الاحتمال الصغير نرتكب خطأ من النوع الثاني (II) وهو غير محتمل إلى أبعد حد . ذلك لأننا سنرفض اللقاح الجديد في الوقت الذي نعلم فيه أن فعاليته بعد مرور عامين على استخدامه هي 60% . ونلاحظ أنه عندما يقترب الفرض البديل من الواحد  $H_1: p = 1$  ، فإن  $\beta$  تقترب من الصفر .

ملاحظات هامة :

إن عملية إنقاص  $\beta$  ممكن دوماً ، وذلك بزيادة حجم المنطقة الحرجة لنرى على سبيل المثال ماذا يحدث لكل من  $\alpha$  و  $\beta$  عندما نغير نقطتنا الحرجة إلى 6.5 . ولنختبر فرض العدم  $P = \frac{1}{5}$  ضد الفرض البديل  $P = 0.6$  .

$$\alpha = \sum_{x=6}^{15} b \left( x; 15, \frac{1}{5} \right) \quad \text{نلاحظ أن :}$$

$$= 1 - \sum_{x=0}^5 b \left( x; 15, \frac{1}{5} \right)$$

$$= 1 - 0.9389$$

$$= 0.0611$$

وأن :

$$\beta = \sum_{x=0}^3 b(x; 15, 0.6) .$$

$$= 0.0338$$

وهكذا نلاحظ أننا انقصنا احتمال ارتكاب خطأ من النوع الثاني على حساب زيادة احتمال ارتكاب الخطأ من النوع الأول . لحسن الحظ أنه يمكن إنقاص كلا الاحتمالين بزيادة حجم العينة . لتعيد نفس التجربة وذلك من أجل  $n = 75$  شخص . لنفرض أننا نرفض فرض العدم  $p = \frac{1}{5}$  فيما إذا اجتاز 25 شخصا أو أكثر العاملين دون التقاطهم الفيروس ، ونقبل عندئذ الفرض البديل  $p > \frac{1}{5}$  . إن النقطة الحرجة في هذه الحالة هي 24.5 ، وجميع النقاط المحرزة والواقعة أعلى هذه النقطة تؤلف ما سميناه بالمنطقة الحرجة . كما أن جميع النقاط المحرزة التي تقع أدناها تشكل منطقة القبول .

سنستخدم المنحنى الطبيعي لحساب احتمال ارتكاب خطأ من النوع الأول (I) . وفي هذه الحالة نلاحظ أن :

$$\mu = np = (75) \left( \frac{1}{5} \right) = 15$$

$$\sigma = \sqrt{npq} = \sqrt{(75) \left( \frac{1}{5} \right) \left( \frac{4}{5} \right)} = 3.464$$

ويوضح القسم المظلل على الشكل (١,٧) هذا الاحتمال .

$\alpha = P$  ( ارتكاب خطأ من النوع الأول )

( فرض العدم صحيح |  $X > 24.5$  )

والقيمة Z المقابلة لـ  $x = 24.5$  هي :

$$Z = \frac{24.5 - 15}{3.464} = 2.742$$

لذلك فإن :

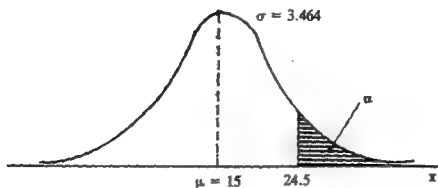
$$\alpha = P(Z > 2.742)$$

$$= 1 - P(Z < 2.742)$$

$$= 1 - 0.9969$$

$$= 0.0031$$

يمكن أن نحدد احتمال ارتكاب خطأ من النوع الثاني ، وذلك باستخدام المنحنى الطبيعي فيما إذا كان فرض العدم خاطئاً والبديل صحيح أى أن  $P = 0.6$  نلاحظ أن :



الشكل (٧، ١)

$$\mu = np = (75)(0.6) = 45$$

$$\sigma = \sqrt{npq} = \sqrt{(75)(0.6)(0.4)} = 4.242$$

ويمثل القسم المظلل من الشكل (٧، ٢) احتمال الوقوع ضمن منطقة القبول عندما يكون  $H_1$  صحيحاً لذلك فإن :

$\beta$  = (ارتكاب خطأً من النوع الثاني)  $= P$

$$= P(X < 24.5 | H \text{ صحيحاً})$$

وتكون قيمة  $Z$  المقابلة للقيمة  $x = 24.5$  هي :

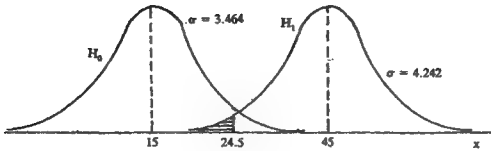
$$Z = \frac{24.5 - 45}{4.242} = -4.83$$

لذلك فإن :

$$\beta = P(Z < -4.83)$$

$$\approx 0$$

والنتائج السابقة تشير إلى أنه قلما يحدث خطأً من النوع الأول أو النوع الثاني عندما تضم التجربة ٧٥ شخصاً .



الشكل (٧، ٢)

مثال (٧، ١)

أوجد كلا من  $\alpha$  و  $\beta$  في الاختبار :

$$H_0 : P = \frac{1}{4}$$

$$H_1 : P = \frac{1}{2}$$

وذلك عندما يكون للإحصاء  $X$  توزيعاً حدائياً في تجربة كررت  $n = 10$  مرات ،  
وبفرض أن المنطقة المخرجة معرفة بالعلاقة  $X \geq 4$  .

الحل

إن احتمال الرفض الخاطئ  $\alpha$  هو :

$$\begin{aligned} \alpha &= P \left[ X \geq 4 \mid P = \frac{1}{4} \right] \\ &= \sum_{x=4}^{10} \binom{10}{x} \left( \frac{1}{4} \right)^x \left( \frac{3}{4} \right)^{10-x} \\ &= 0.2241 \end{aligned}$$



كما أن احتمال القبول الخاطئ  $\beta$  فيحسب بالعلاقة :

$$\begin{aligned}\beta &= P \left( X < 4 \mid P = \frac{1}{2} \right) \\ &= \sum_{x=0}^3 \binom{10}{x} \left( \frac{1}{2} \right)^x \left( \frac{3}{4} \right)^{10-x} \\ &= 0.1719\end{aligned}$$

مثال (٧,٢)

يفرض أن  $\alpha = 0.05$  في المثال (٧,١) برهن أن المنطقة المخرجة هي  $x \geq 6$ .

الحل

نلاحظ أن :

$$\begin{aligned}\alpha = 0.05 &= P \left[ X \geq y \mid P = \frac{1}{4} \right] \\ &= \sum_{x=y}^{10} \binom{10}{x} \left( \frac{1}{4} \right)^x \left( \frac{3}{4} \right)^{10-x}\end{aligned}$$

ونلاحظ في المثال (٧,١) أنه إذا كانت  $y = 4$  لكانت  $\alpha = 0.2281$  وبشكل مشابه إذا كانت  $y = 5$  لكانت  $\alpha = 0.0781$  ، وإذا كانت  $y = 6$  لكانت  $\alpha = 0.0197$  . وبما أن  $y$  عدد صحيح ، لذلك ليس هناك قيمة لـ  $y$  تعطي قيمة لـ  $\alpha = 0.05$  بالضبط ، فإذا كانت  $y = 5$  عندئذ تكون  $\alpha$  كبيرة جداً . أما إذا كانت  $y = 6$  لكانت  $\alpha$  صغيرة جداً . لذلك إذا أخذنا  $y = 6$  فإننا نجد أن  $\alpha$  لن تكون أكبر من 0.05 .

مثال (٧,٣)

لنفرض أننا نرغب في تقدير وسط الإنتاج اليومي في مصنع ، ولنفرض أننا سجلنا الإنتاج اليومي لفترة  $n = 50$  يوماً فكان وسط هذه العينة وانحرافها المعياري بالأطنان كالتالي :

$$\bar{x} = 871, S = 21$$

لنرمز بـ  $\mu$  لوسط الإنتاج اليومي في هذا المصنع . من المعلوم أن  $\bar{x} = 871$  طناً يومياً هو التقدير الأفضل للوسط  $\mu$  . لنختبر الفرض بأن معدل الإنتاج اليومي هو  $\mu = 880$  يومياً ، والذي سنعتبره فرض عدم سنرمز له بالرمز  $H_0$  أى أن :

$$H_0 : \mu = 880$$

ضد الفرض البديل بأن  $\mu = 880$  طناً يومياً أى أن :

$$H_1 : \mu \neq 880$$

نلاحظ أن العينة المستخدمة  $n = 50$  قياساً تخوى على  $\bar{x} = 871, S = 21$  ، وكما أوضحنا فإن الإحصاء  $\bar{x}$  يمثل تقديراً نقطياً لـ  $\mu$  أى أن إحصاء الاختبار هو :

$$Z = \frac{\bar{X} - H_0}{\sigma_{\bar{x}}} = \frac{\bar{X} - H_0}{\sigma / \sqrt{n}}$$

وباستخدام  $S$  كتقريب لـ  $\sigma$  نجد أن :

$$Z = \frac{871 - 880}{21 / \sqrt{50}} = - 3.03$$

ومن أجل  $\alpha = 0.05$  تكون منطقة الرفض  $Z > 1.96$  أو  $Z < - 1.96$  . وبما أن القيمة المحسوبة لـ  $Z$  تقع ضمن منطقة الرفض لذلك نرفض فرض العدم ونستنتج أن  $\mu < 880$  .

#### مثال (٧، ٤)

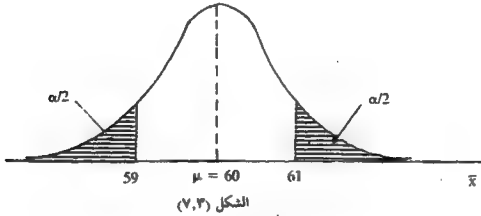
لاختبار فرض العدم ، وهو أن وسط أوزان الطلاب في كلية الهندسة هو 68 كيلوغراماً ضد الفرض البديل في أن هذا الوسط لا يساوى 68 كيلوغراماً أى أن :

$$H_0 : \mu = 60$$

$$H_1 : \mu \neq 60$$

وذلك بفرض أن الانحراف المعياري لمجتمع أوزان الطلاب في هذه الكلية هو  $\sigma = 3$  ، فإننا نختار عينة من طلاب كلية الهندسة بصورة عشوائية حجمها  $n = 32$  طالباً . ثم نحدد قبل كل شيء الإحصاء  $\bar{x}$  والذي سنبنى عليه قرارنا . نلاحظ أن الإحصاء  $\bar{x}$  يمثل أفضل تقدير للوسط  $\mu$  ، وأن لهذا الإحصاء توزيع معانية قريباً من التوزيع الطبيعي بالانحراف المعياري  $\sigma_{\bar{x}} = \sigma / \sqrt{n} = 0.53$  .

لنحدد بعد ذلك النقطة الحرجة والمنطقة الحرجة ومنطقة القبول . نلاحظ أن وسط العينة  $\bar{X}$  يقع بالقرب من القيمة المختبرة 60 . لذلك نأخذ العددين 59 , 61 ثم نعتبر أن المنطقة الحرجة هي المنطقة  $\bar{X} < 59$  و  $\bar{X} > 61$  . انظر الشكل (٧,٣) .



لذلك فإن منطقة القبول ستكون المنطقة  $59 < \bar{X} < 61$  . وهكذا نستنتج أنه إذا وقع وسط العينة  $\bar{X}$  داخل القسم المظلل ( المنطقة الحرجة ) فإننا نرفض  $H_0$  ، أما إذا وقع في المجال ( 59 , 61 ) فإننا نقبلها .

لنحسب احتمال ارتكاب خطأ من النوع الأول ، نعلم أن :

$$\alpha = P ( \bar{X} < 59 | \text{صحيح } H_0 ) + P ( \bar{X} > 61 | \text{صحيح } H_0 )$$

وقيم  $Z$  الموافقة للقيمتين  $\bar{x}_2 = 61$  ,  $\bar{x}_1 = 59$  عندما يكون  $H_0$  صحيحاً هي القيم :

$$Z_1 \frac{59 - 60}{0.53} = -1.886$$

$$Z_2 \frac{61 - 60}{0.53} = 1.886$$

لذلك فإن :

$$\alpha = P ( Z < -1.886 ) + P ( Z > 1.886 )$$

$$\alpha = 2P ( Z < -1.886 )$$

ومنه :

$$\alpha = 0.0588$$

هذا يعني أن 0.06% من العينات ذات الحجم  $n = 30$  مستقودنا إلى رفض الفرض  $\mu = 60$  كيلوغراما في الوقت التي تكون فيه هذه الفرض صحيحة . نلاحظ أنه لإنقاص  $\alpha$  ، علينا أن نزيد حجم العينة ، أو أن نجعل منطقة القبول أوسع مما هي عليه الآن . لنفرض مثلا أن  $n = 50$  نلاحظ أن :

$$\sigma_{\bar{x}} = 30 / \sqrt{50} = 0.424$$

كما أن :

$$Z_1 = \frac{59 - 60}{0.424} = - 2.358$$

$$Z_2 = \frac{61 - 60}{0.424} = 2.358$$

لذلك فإن :

$$\begin{aligned}\alpha &= P(Z < - 2.358) + P(Z > 2.358) \\ &= 2P(Z < - 2.358)\end{aligned}$$

ومنه :

$$= 0.0182$$

ولا يكفي إنقاص قيمة  $\alpha$  لكي نضمن إجراءات الاختبار الذي نجربه جيدة . وعلينا تحديد قيمة  $\beta$  من أجل مختلف الفروض البديلة ، والتي نشعر أنها ستكون مقبولة إذا كانت صحيحة . لذلك فإن من الضروري أن نرفض  $H_0$  عندما يكون الوسط مساويا لإحدى القيم  $62 \leq \mu$  أو  $58 \leq \mu$  . وفي هذه الحالة علينا أن نحدد احتمال ارتكاب خطأ من النوع الثاني  $\beta$  وفحصه ، وذلك من أجل الفرضين  $\mu = 62$  ،  $\mu = 58$  .

ويكفي فقط ( وذلك بسبب التناظر ) أن نفكر في احتمال قبول فرض العدم  $\mu = 60$  ، وذلك عندما يكون الفرض البديل  $\mu = 62$  صحيحاً . فإذا وقع وسط العينة  $\bar{x}$  بين 59 و 61 عندما يكون  $H_1$  صحيحاً . فإننا نكون قد ارتكبنا خطأ من النوع II . بالعودة إلى الشكل (٤، ٧) نجد أن :

$$\beta = P(59 < \bar{X} < 61 | \text{صحيحاً } H_1)$$

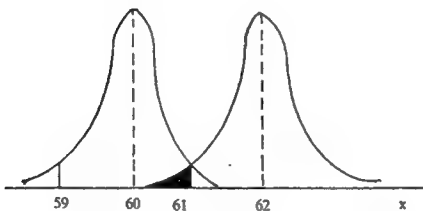
نلاحظ أن قيم Z المقابلة لـ  $\bar{x}_1 = 59$  ،  $\bar{x}_2 = 61$  عندما يكون  $H_1$  صحيحاً هي :

$$Z_1 = \frac{59 - 62}{0.424} = -7.07$$

$$Z_2 = \frac{61 - 62}{0.424} = -2.36$$

لذلك فإن :

$$\begin{aligned}\beta &= P(-7.07 < Z < -2.36) \\ &= P(Z < -2.36) - P(Z < -7.07) \\ &\approx 0.0091 - 0.0000 \\ &= 0.0091\end{aligned}$$



الشكل (٤، ٧)

إذا كانت القيمة الحقيقية لـ  $\mu$  هي القيمة البديلة  $\mu = 58$  ، فعندئذ ستكون قيمة  $\beta$  هي نفس العدد 0.0132 . من أجل جميع القيم  $\mu < 58$  أو  $\mu > 62$  ، فإن قيمة  $\beta$  ستكون صغيرة وذلك من أجل  $n = 58$  . لذلك فإنه سيطرأ تغير طفيف حول قبول  $H_0$  في الوقت الذي يكون فيه خاطئاً .

إن احتمال ارتكاب خطأ من النوع II سيزداد بسرعة عندما تقترب القيمة الحقيقية لـ  $\mu$  من القيمة المفروضة دون أن تساويها ، وهذه الحالة عادية بالطبع وذلك عندما لا نفكر في ارتكاب خطأ من النوع II فمثلاً إذا كان الفرض البديل  $\mu = 60.5$  صحيحاً ، فإننا لا نفكر في ارتكاب خطأ من النوع II وذلك باستنتاج أن الإجابة الصحيحة هي  $\mu$

$\mu = 60$ ، واحتمال حصول مثل هذا الخطأ سيكون عاليا جدا عندما تكون  $n = 58$ . بالعودة إلى الشكل (٧،٥) نجد أن :

( عندما يكون  $H_1$  صحيحاً  $| 59 < \bar{x} < 61$  )

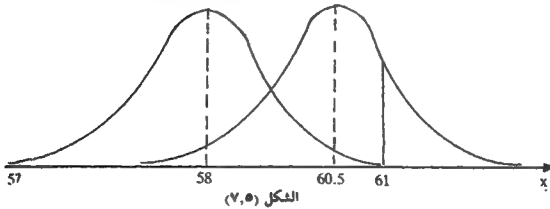
وقيم  $Z$  المقابلة للقيم  $\bar{x}_1 = 59$  و  $\bar{x}_2 = 61$  من أجل  $\mu = 60.5$

$$Z_1 = \frac{59 - 60.5}{0.424} = -3.54$$

$$Z_2 = \frac{61 - 60.5}{0.424} = 1.18$$

لذلك فإن :

$$\begin{aligned}\beta &= P(-3.54 < Z < -1.18) \\ &= P(Z < 1.18) - P(Z < -3.54) \\ &= 0.8810 - 0.0000 \\ &= 0.8810\end{aligned}$$



توضيح الأمثلة السابقة الخواص الهامة التالية :

- (١) الخطأ من النوع I والخطأ من النوع II مرتبطان ببعضهما ، فأى زيادة في احتمال أحدهما يؤدي إلى نقصان في احتمال الآخر .
- (٢) إن حجم المنطقة المخرجة ، واحتمال ارتكاب خطأ من النوع الأول يمكن

انقاصهما بتعديل القيمة الحرجة .

(٣) إن أى زيادة فى حجم العينة  $n$  سينقص كلا من  $\alpha$  و  $\beta$  فى وقت واحد .

(٤) إذا كان فرض العدم خاطئاً ، فعندئذ يكون  $\beta$  أعظمياً عندما تكون القيمة الحقيقية للمتغير قريبة ( مجاورة ) للقيمة المختبرة . تقابل المسافة الكبيرة بين القيمة الحقيقية والقيمة المختبرة قيمة صغيرة لـ  $\beta$  .

### (٧,٣) الاختبارات وحيدة وثنائية الذيل One-tailed and two-tailed tests

نسمى اختبار فرض إحصائى يكون فيه فرض البديل ذو طرف واحد مثل :

$$\begin{aligned} H_0 : \theta &= \theta_0 \\ H_1 : \theta &> \theta_0 \end{aligned} \quad \text{أو :}$$

$$\begin{aligned} H_0 : \theta &= \theta_0 \\ H_1 : \theta &< \theta_0 \end{aligned}$$

باختبار وحيد الذيل . وتكون المنطقة الحرجة للفرض البديل  $\theta > \theta_0$  تماماً فى الذيل الأيمن للتوزيع ، بينما تكون المنطقة الحرجة للفرض البديل  $\theta < \theta_0$  فى الذيل الأيسر تماماً . يوضح المثال (٧,٣) الاختبار الوحيد الذيل .

نسسمى الاختبار الذى يكون فيه الفرض البديل ذو طرفين مثل :

$$\begin{aligned} H_0 : \theta &= \theta_0 \\ H_1 : \theta &\neq \theta_0 \end{aligned}$$

باختبار ثنائى الذيل . تؤلف قيم الفرض البديل فى الذيلين ما يسمى بالمنطقة الحرجة . وكمثال على الاختبار ثنائى الذيل نذكر المثال (٧,٤) الذى سقناه حول اختبار وسط أوزان طلاب كلية الهندسة . وكنا قد فرضنا أن فرض العدم هو  $\mu = 60$  كيلوغراما . كما أن الفرض البديل ثنائى الذيل هو  $\mu \neq 60$  .

### ملاحظة هامة

عند اختبار فرض يتعلق بمجتمع منقطع فإننا نقوم بشكلٍ كفي باختبار المنطقة الحرجة وتحديد حجمها . فإذا كان حجم هذه المنطقة  $\alpha$  كبيراً جداً ، فإن بإمكاننا

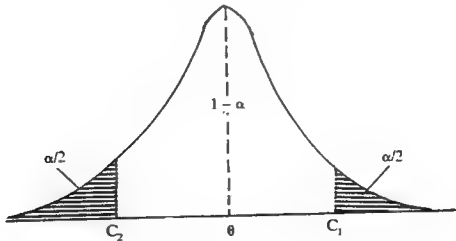
انقاصه بتعديل موضع القيمة الحرجة . أما عندما نريد اختبار فرضية حول مجتمع مستمر ، فإننا نقوم باختبار قيمة  $\alpha$  لتكون مساوية لـ 0.05 أو 0.01 على الغالب ، وبعدها نجد القيمة الحرجة C الموافقة . فمثلاً في الاختبار التثنائي الذيل بمستوى المعنوية  $\alpha = 0.05$  من الأهمية . فإن القيمتين الحرجتين للإحصاء ذى التوزيع الطبيعي المعيارى ستكونان - 1.96 و  $Z_{0.025} = 1.96$  ، وإذا استخدمنا مصطلح قيم Z فستكونان عندئذ المنطقة الحرجة من الشكل  $Z < -1.96$  و  $Z > 1.96$  . أما منطقة القبول فهي  $-1.96 < Z < 1.96$  .

نقول بأن اختبارنا فعالاً إذا رفضنا  $H_0$  تحت مستوى من المعنوية قدره  $\alpha = 0.05$  ، ويعتبر هذا الاختبار أكثر فعالية إذا تم رفض  $H_0$  تحت مستوى من المعنوية قدره  $\alpha = 0.01$  . أما إذا كنا نرفض  $H_0$  مجرد ألا يكون مساوياً لـ  $\theta$  أى سواء كان  $\theta_0$  أكبر أم أصغر من  $\theta$  فعندما نعبّر عن فرض العدم والبديل بالشكل التالى :

$$H_0 : \theta = \theta_0$$

$$H_1 : \theta \neq \theta_0$$

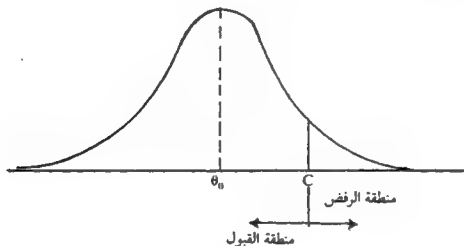
ويمكن القول بأن أفضل اختبار يتوفر لنا عملياً هو الاختبار الذى تتوضع فيه منطقة الرفض على التساوى في ذيلى المنحنى . نلاحظ أن اختبارنا السابق هو ثنائى الذيل . وعندما نقتطع مساحة تساوى  $\frac{\alpha}{2}$  من كل ذيل من ذيلى المنحنى ، فإننا نقوم عندئذ بتحديد نقطتين  $C_1$  و  $C_2$  بحيث يكون  $H_0 = \frac{\alpha}{2}$  صحيح  $P(\hat{\theta} \geq C_1 | H_0 = \frac{\alpha}{2})$  و  $P(\hat{\theta} \leq C_2 | H_0 = \frac{\alpha}{2})$  صحيح . هذا ونرفض الفرض  $H_0$  إذا كانت قيمة  $\hat{\theta}$  إلى يمين  $C_1$  أو إلى يسار  $C_2$  ونقبله فيما عدا ذلك . يوضح الشكل (٧،٦) شكل منطقة الرفض وتوزيع  $\hat{\theta}$  عندما يكون  $H_0$  صحيحاً .



الشكل (٧،٦)



لنفرض أننا نرغب في اختبار فرض يتعلق بوسيط  $\theta$  وأن التقدير النقطة  $\hat{\theta}$  لهذا الوسيط له توزيع طبيعي بوسط يساوى  $\theta_0$  وانحراف معياري  $\sigma_{\hat{\theta}}$ . فإذا كان فرض العدم  $H_0: \theta = \theta_0$  فعندئذ سيتوزع التقدير  $\hat{\theta}$  وفقا للتوزيع الطبيعي بوسط قدره  $\theta_0$  ، وذلك كما هو واضح على الشكل (٧,٧) .



الشكل (٧,٧)

لنفرض أننا نهم برفض الفرض  $H_0$  عندما يكون  $\theta > \theta_0$  ، فعندئذ يكون الفرض البديل  $H_1: \theta > \theta_0$  ويمكن البرهان على أن أفضل اختبار هو ذلك الذى يرفض  $H_0$  عندما يكون  $\hat{\theta}$  كبير . والمقصود بكبير  $\hat{\theta}$  هو أن تكون المسافة بينه وبين  $\theta_0$  مساوية لعدد من الانحرافات المعيارية  $\sigma_{\hat{\theta}}$  . يوضح الشكل (٧,٧) موقع منطقتى الرفض والقبول . نسمى القيمة C التى تفصل منطقة الرفض عن منطقة القبول بالقيمة الحرجة لإحصاء الاختبار . هذا ويحدد مستوى المعنوية الاختبار  $\alpha$  موقع C لأن احتمال أن يكون  $\hat{\theta} \geq C$  علما أن  $H_0$  صحيح ( أى احتمال رفض  $H_0$  على أنه صحيح ) يساوى  $\alpha$  ونعبر عن ذلك بالعلاقة  $P(\hat{\theta} \geq C | H_0) = \alpha$  . وعلى سبيل المثال إذا كان  $\alpha = 0.05$  ، فإن  $C = \theta_0 + 1.645 \sigma_{\hat{\theta}}$  . ونلاحظ أن منطقة الرفض تقع في الذيل الأيمن للمنحنى ، لذلك فاختبارنا هذا هو اختبار وحيد الذيل .

#### (٧,٤) اختبار وسط وتباين مجتمع إحصائي Tests concerning means and variances

نفرض أن  $\mu$  يمثل وسط مجتمع إحصائي تباينه  $\sigma^2$  معلوم . لنختبر فرض العدم  $H_0: \mu = \mu_0$  ضد الفرض البديل ثنائي الذيل  $H_1: \mu \neq \mu_0$  . وك تقدير مناسب لهذا

الوسيط نأخذ الإحصاء  $\bar{X}$  ونبنى عليه قرارنا . من المعلوم كما ورد في الفصل الخامس أن للمتغير  $\bar{X}$  توزيعاً قريباً من التوزيع الطبيعي بالوسيط :  $\mu_{\bar{X}} = \mu$  والتباين  $\sigma_{\bar{X}}^2 = \frac{\sigma^2}{n}$  ( يمثل هذا العدد  $n$  حجم عينة مختارة بصورة عشوائية من المجتمع المدروس ) . إذا استخدمنا كمستوى معنوية لهذا الاختبار العدد  $\alpha$  ، فعندئذ يمكن أن نجد عددين موجبين  $\bar{x}_1$  و  $\bar{x}_2$  بحيث يكون  $\bar{x}_1 < \bar{X} < \bar{x}_2$  يحددان منطقة القبول . ويؤلف ذيل التوزيع  $\bar{X} < \bar{x}_1$  و  $\bar{X} > \bar{x}_2$  ما يسمى بالمنطقة الحرجة .

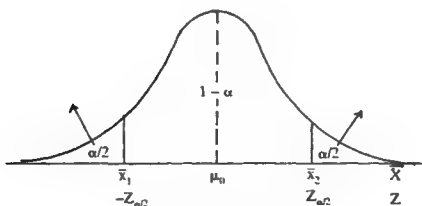
يمكن أن نعرف المنطقة الحرجة باستخدام قيم  $Z$  ، وذلك باستخدام علاقة التحويل :

$$Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}$$

لذلك ومن أجل مستوى المعنوية  $\alpha$  ، فإنه يمكن مشاهدة القيم الحرجة للمتغير  $Z$  والموافقة للقيم  $\bar{x}_1$  و  $\bar{x}_2$  على الشكل (٧،٨) ، حيث إن :

$$-Z_{\alpha/2} = \frac{\bar{x}_1 - \mu_0}{\sigma / \sqrt{n}}$$

$$Z_{\alpha/2} = \frac{\bar{x}_2 - \mu_0}{\sigma / \sqrt{n}}$$



الشكل (٧،٨)

فإذا وقعت  $\bar{x}$  في منطقة القبول  $\bar{x}_1 < \bar{x} < \bar{x}_2$  ، فنعتقد ستقع  $\bar{x} - \mu_0$  ضمن المنطقة  $-Z_{\alpha/2} < Z < Z_{\alpha/2}$  . نستنتج أن  $\mu = \mu_0$  . وهذا يعني أننا سنرفض  $H_0$  ونقبل الفرض البديل  $\mu \neq \mu_0$  .

إن إجراءات الاختبار التي درسناها سابقا تكافئ إيجاد  $100\%(1 - \alpha)$  مجال ثقة للمتوسط  $\mu$  وقبول  $H_0$  إذا كانت  $H_0$  متممة للمجال السابق . أما إذا رفضت  $\mu_0$  خارج المجال فإننا سنرفض الفرض  $H_0$  لصالح الفرض البديل  $H_1$  . وهكذا فإنه عندما يضع الدارس استنتاجات ما حول وسط  $\mu$  لمجتمع تباينه  $\sigma^2$  معلوم سواء أكانت بإنشاء مجال ثقة أو عن طريق اختبار الفرضيات ، فإنه يستخدم نفس الإحصاء .

$$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$$

وبشكل عام إذا استخدم الدارس إحصاء ما لإنشاء مجال ثقة للمتغير  $\theta$  سواءا أكان الإحصاء  $Z$  ،  $T$  ،  $\chi^2$  أو  $F$  ، فإن بإمكان استخدام نفس الإحصاء لاختبار الفرض في أن المتغير ( الوسيط )  $\theta$  يساوى قيمة محددة  $\theta_0$  ضد أى فرض بديل . وبالطبع فإننا سنطبق جميع الافتراضات الأساسية التي وردت في الفصل السادس والمتعلقة باستخدام إحصاء ما ، وذلك لوصف الاختبارات المعروضة في هذا الفصل . ويعنى هذا بشكل أساسي أننا نسحب عيناتنا من مجتمعات لها توزيعات قريبة من التوزيع الطبيعي . ومع ذلك فإن الإحصاء  $Z$  يمكن استخدامه لاختبار فرض حول وسطاء مجتمعات غير طبيعية فيها  $n \geq 30$  .

ويوضح الجدول (٧،٢) الطرق المستخدمة لاختبار فرض معين  $H_0$  . كما يعطى المناطق الحرجة المناسبة للفروض وحيدة أو ثنائية الذيل . يمكن أن نجمل الخطوات المتبعة في اختبار فرض متعلق بوسيط مجتمع  $\theta$  ضد فرض بديل بالخطوات التالية :

نحدد أولا فرض العدم  $\theta = \theta_0$  :  $H_0$  ، وبعد ذلك نحدد الفرض البديل  $\theta \neq \theta_0$  أو  $\theta > \theta_0$  أو  $\theta < \theta_0$  :  $H_1$  ثم نختار مستوى المعنوية  $\alpha$  ونحدد الاختبار الإحصائي المناسب ، وبعد ذلك المنطقة الحرجة . نحسب أيضا قيمة إحصاء الاختبار ، وذلك من خلال عينة حجمها  $n$  . وأخيرا فإننا نتخذ قراراً حول رفض  $H_0$  إذا وقعت قيمة الإحصاء المحسوب في المنطقة الحرجة ونقبلها في عكس ذلك .

$H_0$	إحصاء الاختبار	$H_1$	المنطقة الحرجة
$\mu = \mu_0$	$Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}$	$\mu < \mu_0$	$Z < -Z_\alpha$
		$\mu > \mu_0$	$Z > Z_\alpha$
		$\mu \neq \mu_0$	$Z < -Z_{\alpha/2}$ $Z > Z_{\alpha/2}$
$\mu = \mu_0$	$T = \frac{\bar{X} - \mu_0}{S / \sqrt{n}} ; v = n - 1$	$\mu < \mu_0$	$T < -t_\alpha$
		$\mu > \mu_0$	$T > t_\alpha$
		$\mu \neq \mu_0$	$T < -t_{\alpha/2}$ $T > t_{\alpha/2}$
$\mu_1 - \mu_2 = d_0$	$Z = \frac{(\bar{X}_1 - \bar{X}_2) - d_0}{\sqrt{(\sigma_1^2/n_1) + (\sigma_2^2/n_2)}}$	$\mu_1 - \mu_2 < d_0$	$Z < -Z_\alpha$
		$\mu_1 - \mu_2 > d_0$	$Z > Z_\alpha$
		$\mu_1 - \mu_2 \neq d_0$	$Z < -Z_{\alpha/2}$ $Z > Z_{\alpha/2}$
$\mu_1 - \mu_2 = d_0$	$T = \frac{(\bar{X}_1 - \bar{X}_2) - d_0}{S_p \sqrt{(1/n_1) + (1/n_2)}}$ $v = n_1 + n_2 - 2, \sigma_1 = \sigma_2$ غير أنهما غير معلومين $S_p^2 = \frac{(n_1 - 1) S_1^2 + (n_2 - 1) S_2^2}{n_1 + n_2 - 2}$	$\mu_1 - \mu_2 < d_0$	$T < -t_\alpha$
		$\mu_1 - \mu_2 > d_0$	$T > t_\alpha$
		$\mu_1 - \mu_2 \neq d_0$	$T < -t_{\alpha/2}$ $T > t_{\alpha/2}$
$\mu_1 - \mu_2 = d_0$	$T' = \frac{(\bar{X}_1 - \bar{X}_2) - d_0}{\sqrt{(S_1^2/n_1) + (S_2^2/n_2)}}$	$\mu_1 - \mu_2 < d_0$	$T' < -t_\alpha$
		$\mu_1 - \mu_2 > d_0$	$T' > t_\alpha$
		$\mu_1 - \mu_2 \neq d_0$	$T' < -t_{\alpha/2}$ $T' > t_{\alpha/2}$

$H_0$	إحصاء الاختبار	$H_1$	المنطقة الحرجة
	$S^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$ $(S_1^2/n_1 + S_2^2/n_2)^2$ $v = \frac{(S_1^2/n_1)^2}{n_1 - 1} + \frac{(S_2^2/n_2)^2}{n_2 - 1}$ <p>وغير معلومين <math>\sigma_1</math> <math>\sigma_2</math></p> $\chi^2 = \frac{(n - 1) S^2}{\sigma_0^2}$		
$\sigma^2 = \sigma_0^2$		$\sigma^2 < \sigma_1^2$ $\sigma^2 > \sigma_0^2$ $\sigma^2 \neq \sigma_0^2$	$\chi^2 < \chi_1^2 - \alpha$ $\chi^2 > \chi_\alpha^2$ $\chi^2 < \chi_1^2 - \alpha/2$ $\chi^2 > \chi_{\alpha/2}^2$
$\sigma_1^2 = \sigma_2^2$	$F = S_1^2/S_2^2$ $v_1 = n_1 - 1$ $v_2 = n_2 - 1$	$\sigma_1^2 < \sigma_2^2$ $\sigma_1^2 > \sigma_2^2$ $\sigma_1^2 \neq \sigma_2^2$	$F < f_1 - \alpha$ ( $v_1, v_2$ ) $F > f_\alpha$ ( $v_1, v_2$ ) $F < f_1 - \alpha/2$ ( $v_1, v_2$ ) $F > f_{\alpha/2}$ ( $v_1, v_2$ )

الجدول (٧, ٢)

## مثال (٧, ٥)

من المعلوم أن واحداً من عشرة من المدخنين على وجه التقريب في المملكة العربية السعودية يفضلون نوعاً معيناً من السجائر . وقد قامت الشركة المنتجة لهذا النوع بحملة دعائية واسعة لهذا النوع داخل المملكة وبعدها ، ويقصد اختبار فعالية هذه الحملة قامت الشركة بأخذ عينة مؤلفة من 200 مدخن فتبين لها أن 26 مدخناً منهم يفضلون هذا النوع . فهل كانت هذه الحملة ذات فائدة ؟

## الحل

نلاحظ أن التجربة تحقق شروط التجربة الحدائية . نفرض أن  $p$  يمثل احتمال تفضيل

مدخن ما في المملكة للنوع المذكور من السجائر ، ولنتخبر فرض العدم .

$$H_0 : p = 0.10$$

ضد الفرض البديل :

$$H_1 : p > 0.10$$

ومن الطبيعي أن تكون  $p$  كذلك ، لأننا نهم بالكشف عن أية زيادة في الاحتمال  $p$  منطلقين من الاعتقاد بأن الحملة الدعائية إن لم تكن مفيدة فهي على أى حال لا يمكن أن تكون ضارة فتخفض من قيمة  $p$  . هذا وأن أفضل اختبار لمثل هذا الفرض هو الاختبار الناتج عن وضع كامل منطقة الرفض في الذيل الأيمن لتوزيع إحصاء الاختبار . نعلم أن

$$Z = \frac{\hat{p} - p_0}{\sigma_{\hat{p}}} = \frac{p - p_0}{\sqrt{p_0 q_0 / n}}$$
 ويكون إحصاء الاختبار عندئذ  $\hat{p} = \frac{x}{n}$  وهذا يعنى أن  $Z = \frac{x - np_0}{\sqrt{np_0 q_0}}$  ، حيث فرضنا أن  $P = P_0$  في عبارة الانحراف المعياري لـ  $p$  وهي  $\sqrt{npq}$  . وإذا اخترنا مستوى المعنوية  $\alpha = 0.05$  ، فإننا نلاحظ من الجدول IV أن  $Z = 1.645$  . وعلى هذا فإننا سنرفض  $H_0$  عندما يكون  $Z > 1.645$  . وبحساب  $Z$  نجد أن :

$$Z = \frac{26 - 20}{\sqrt{200(0.1)(0.9)}} = 1.41$$

والقيمة الأخيرة للإحصاء لا تقع ضمن منطقة الرفض ، وبالتالي فإننا نقبل الفرض  $H_0$  . وهذا يعنى أن الحملة الدعائية كانت مفيدة حقاً .

### مثال (٧,٩)

يفرض أن وسط الزمن اللازم لتسجيل أى طالب بالنسبة لجميع الصفوف في كلية الهندسة هو 50 دقيقة وأن الانحراف المعياري هو 10 دقائق . استخدمنا طريقة حديثة أخرى للتسجيل ثم سحبنا عينة من الطلاب بصورة عشوائية مؤلفة من اثني عشر طالباً ، فوجد أن لها وسط تسجيل قدره 42 دقيقة وانحرافاً معيارياً قدره 11.9 دقيقة . اختبر الفرض في أن وسط المجتمع أقل من 50 ، وذلك باستخدام ١ - مستوى معنوية  $\alpha = 0.05$  - ٢ - ثم باستخدام مستوى معنوية  $\alpha = 0.01$  . افرض أن مجتمع الأزمنة طبيعي .

### الحل

(١) نلاحظ أن فرض العدم هو دقيقة 50  $H_0 : \mu = 50$

- (٢) كما نلاحظ أن الفرض البديل هو دقيقة  $50 < \mu$  :  $H_1$
- (٣) وأن مستوى المعنوية في الحالة الأولى  $\alpha = 0.05$  ، وفي الحالة الثانية  $\alpha = 0.01$  .
- (٤) لتحديد المنطقة الحرجة نعود إلى الجدول (٧،٢) ونلاحظ أنه من أجل :

$$T = \frac{\bar{X} - \mu_0}{S / \sqrt{n}} : \nu = 11$$

أن  $T < -1.796$  من أجل مستوى المعنوية  $\alpha = 0.05$  وأن  $T < -2.718$  من أجل  $\alpha = 0.01$

(٥) لنحسب من خلال العينة ذات الحجم  $n = 12$  قيمة الإحصاء  $T$  . نلاحظ أن  $x = 42$  دقيقة ،  $S = 11.9$  و  $n = 12$  لذلك فإن :

$$t = \frac{42 - 50}{11.9 / \sqrt{12}} = -2.33$$

(٦) ثم نتخذ قرارنا التالي : نرفض الفرض  $H_0$  بمستوى المعنوية  $0.05$  ونقبله من أجل مستوى المعنوية  $0.01$  . وهذا يعنى بشكل جوهري أن المرجح أن يكون الوسط الحقيقي لزمن التسجيل أقل من  $50$  دقيقة .

### مثال (٧،٧)

طور مهندسو مصنع للجهيزات الرياضية نوعا من الخيوط لصيد الأسماك ، وادعوا أن لهذا النوع وسط قوة انقطاع قدره  $\mu = 8$  كيلوغراماً ، وانحرافاً معيارياً قدره  $0.5$  كيلوغراماً . اختبر فرض العدم  $H_0 : \mu = 8$  كيلوغراماً ضد الفرض البديل  $H_1 : \mu > 8$  كيلوغراماً ، وذلك بفرض أن اخترنا عينة عشوائية مؤلفة من  $50$  خيطاً . وجربت فأعطت وسطاً عينياً لقوة الانقطاع قدره  $7.5$  كيلوغراماً . استخدم مستوى المعنوية  $\alpha = 0.01$  .

### الحل

نلاحظ أن  $H_0 : \mu = 8$  وأن  $H_1 : \mu \neq 8$  . أما بالنسبة لمستوى المعنوية فهو  $\alpha = 0.01$  . لذلك فإن

المنطقة الحرجة هي  $Z < -2.58$  و  $Z > 2.58$  حيث يمثل  $Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}$  . فإذا فرضنا أن  $\bar{x} = 7.8$  و  $n = 50$  فإننا نجد أن :

$$Z = \frac{7.8 - 8}{0.5 / \sqrt{50}} = -2.828$$

وحيث أن هذه القيمة تقع ضمن منطقة الرفض ، لذلك فإننا نرفض العدم  $H_0$  ، ونستنتج أن وسط قوة الانقطاع ليس مساوياً لـ 8 كيلو غراما ، بل هو أقل من ثمانية كيلو غرامات .

### مثال (٧,٨)

أجريت تجربة لمقارنة تآكل نوعين من المواد . فأخذت عينة مؤلفة من 12 قطعة من المادة الأولى واختبرت بتعريض كل قطعة من قطعها لآلة لقياس التآكل . كما اختبرت عينة أخرى من المادة الثانية مؤلفة من عشر قطع واختبرت بنفس الطريقة ولوحظ عمق التآكل في كل حالة . وقد بلغ وسط التآكل بالنسبة لمادة العينة الأولى 85 وحدة وانحرافها المعياري 4 ، بينما بلغ وسط التآكل بالنسبة لمادة العينة الثانية 81 وانحرافها المعياري 5 . اختبر الفرض التالي : أن للمادتين السابقتين نفس التآكل بمستوى من المعنوية قدره  $\alpha = 0.1$  . وذلك إذا فرضنا أن مجتمعى المادتين السابقتين توزيعاً قريباً من التوزيع الطبيعي بتباينين متساويين .

### الحل

نفرض أن  $\mu_1, \mu_2$  يمثلان وسطى مجتمعى المادتين الأولى والثانية على الترتيب ، لتتبع الخطوات الست التالية :

$$H_0 : \mu_1 = \mu_2 \text{ أو } \mu_1 - \mu_2 = 0 \quad (١)$$

$$H_1 : \mu_1 \neq \mu_2 \text{ أو } \mu_1 - \mu_2 \neq 0 \quad (٢)$$

$$\alpha = 0.1 \quad (٣)$$

(٤) المنطقة الحرجة هي  $T < -1.725$  و  $T > 1.725$  حيث يمثل :

$$T = \frac{(\bar{X}_1 - \bar{X}_2) - d_0}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} ; = 20$$



(٥) بالحساب لدينا  $\bar{x}_2 = 81, S = 5, n_2 = 10, \bar{x}_1 = 85, S = 4, n_1 = 12$  لذلك فإن :

$$S_p = \sqrt{\frac{(11)(16) + (9)(25)}{12 + 10 - 2}} = 4.478$$

$$t = \frac{(85 - 81) - 0}{4.478 \sqrt{(1/12) + (1/10)}} = 2.07$$

(٦) الاستنتاج : بما أن  $t = 2.07$  ( قد وقعت في منطقة الرفض ) لذلت فإننا سنرفض فرض العدم  $H_0$  ونستنتج أنه ليس للمادتين السابقتين نفس التأكل .

### مثال (٧,٩)

استخدمت خمس عينات من مادة حديدية لتحديد ما إذا كان هناك فرقاً بين التحليل الكيميائي المخبري والتحليل الفلوري بأشعة X لمعرفة كمية الحديد الموجودة في هذه المادة . نفرض أن كل عينة قد جزئت إلى عيتين جزئيتين ، وأنه قد طبقت على كل منهما طريقتي التحليل السابقتين . وقد أوضحت نتائج التحليل هذه كمية الحديد الموجودة بالجدول التالي :

### العينة

التحليل	1	2	3	4	5
بأشعة X	2.0	2.0	2.3	2.1	2.4
بالتحليل الكيميائي	2.2	1.9	2.5	2.3	2.4

اختبر بمستوى من المعنوية قدره  $\alpha = 0.05$  ما إذا كانت طريقتا التحليل المذكورتين تعطيان في الوسط نفس النتائج . وذلك بفرض أن المجتمعات طبيعية .

### الحل

نفرض أن وسط كمية الحديد المحددة بالطريقة الكيميائية هي  $\mu_1$  ، وبالطريقة الشعاعية هي  $\mu_2$  . نلاحظ باستخدام الخطوات الست المذكورة سابقاً أن :

$$H_0: \mu_1 = \mu_2, \mu_D \neq 0 \quad (١)$$

$$H_1: \mu_1 \neq \mu_2, \mu_D \neq 0 \quad (٢)$$

$$\alpha = 0.05 \quad (٣)$$

$$(٤) \quad \text{بالعودة إلى السطر السادس في الجدول (٧, ٢) وبفرض أن } T = \frac{\bar{D} - d_0}{S / \sqrt{n}} = 4.2$$

درجات من الحرية ، فإننا نلاحظ أن المنطقة الحرجة هي  $T < -2.776$  و  $T > 2.776$

(٥) الحسابات

التحليل الشعاعى	التحليل الكيميائى	$d_i$	$d_i^2$
2.0	2.2	-0.2	0.04
2.0	1.9	-0.1	0.01
2.3	2.5	-0.2	0.04
2.1	2.3	-0.2	0.04
2.4	2.4	0.0	0.00
		-0.5	0.13

ونجد أن  $\bar{D} = -0.5/5 = -0.1$  وأن :

$$S_d = [(5)(0.13) - (0.5)^2] / (5)(4) = 0.2$$

وبأخذ الجذر التربيعى نجد أن  $S_d = 0.14142$

لذلك فإن :

$$t = \frac{-0.1 - 0}{0.14142 / \sqrt{5}} = -1.6$$

(٦) الاستنتاج : بما أن القيمة  $t = -1.6$  تقع ضمن منطقة القبول ، لذلك فإننا نقبل الفرض  $H_0$  ونستنتج أنه لا يوجد فرق جوهري بين طريقتى التحليل المذكورتين .

مثال (٧, ١٠)

يدعى صاحب للبطاريات أن عمر بطارياته له توزيعاً قريباً من التوزيع الطبيعي بانحراف معيارى قدره 0.9 سنة . أخذت عينة عشوائية حجمها  $n = 10$  من هذه البطاريات ، فوجد أن لها انحرافاً معيارياً عينياً قدره 1.2 سنة . فهل تعتقد أن  $\sigma > 0.9$  سنة ؟ استخدم مستوى المعنوية  $\alpha = 0.05$  .

## الحل

نلاحظ باتباع الخطوات الست المعروفة أن :

$$H_0: \sigma^2 = 0.81 \quad (١)$$

$$H_1: \sigma^2 > 0.81 \quad (٢)$$

$$\alpha = 0.05 \quad (٣)$$

(٤) من السطر السابع في الجدول (٧،٢) نلاحظ وبفرض أن  $\chi^2 = (n - 1) S^2 / \sigma_0^2$  ونلاحظ وبفرض أن  $\chi^2 > 16.919$  حيث إن  $v = 9$  ، أن المنطقة الحرجة في هذه الحالة هي  $\chi^2 > 16.919$  (٥) الحسابات : نلاحظ أن  $S^2 = 1.44$  ،  $n = 10$  وأن :

$$\chi^2 = \frac{(9)(1.44)}{0.81} = 16.0$$

(٦) الاستنتاج : بما أن القيمة  $\chi^2 = 16.0$  تقع ضمن منطقة القبول ، لذلك فإننا نقبل الفرض  $H_0$  ونستنتج أنه لا يوجد أى سبب للشك في أن  $\sigma = 0.9$  سنة .

## (٧،٥) اختبار حجم العينة لاختبار الوسط

## Choice of sample size for testing mean

يمكن للمجرب أن يتحكم بمستوى معنوية اختبار إحصاء الفروض ، بينما يمكن التحكم بـ  $\beta$  أو قوة هذا الاختبار والذي نرمز له بالرمز  $1 - \beta$  باستخدام حجم عينة مناسبة . سنناقش في هذا الفصل طريقة اختبار حجم عينة للاختبارات التي تستلزم وسطاً أو ( وسطين ) لمجمعاتها . نلاحظ أن مسألة تحديد حجم العينة الضروري لتحقيق قوة اختبار معينة هي مسألة سهلة في الحالة التي يكون فيها التوزيع المستخدم طبيعياً وتباينه ( أو تبايناته ) معلوماً . لنفرض أننا نرغب في اختبار الفرض  $H_0: \mu = \mu_0$  وذلك ضد الفرض البديل  $H_1: \mu > \mu_0$  بمستوى من المعنوية قدره  $\alpha$  . وذلك بفرض أن تباين المجتمع المدروس  $\sigma^2$  معلوماً . فعلى سبيل المثال لنفرض أن الفرض البديل المحدد هو  $\mu = \mu_0 + \delta$  عندئذ تعطى قوة هذا الاختبار بالعلاقة التالية :

$$1 - \beta = P \left[ \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}} > Z_{\alpha} | H_1 \right]$$

$$= P \left[ \frac{\bar{X} - (\mu_0 + \delta)}{\sigma / \sqrt{n}} > Z_{\alpha} - \frac{\delta}{\sigma / \sqrt{n}} \mid \mu = \mu_0 + \delta \right]$$

وبالنسبة للفرض البديل  $\mu = \mu_0 + \delta$  ، فإن للإحصاء  $\frac{\bar{X} - (\mu_0 + \delta)}{\sigma / \sqrt{n}}$  توزيعاً طبيعياً معيارياً  
لذلك فإن :

$$1 - \beta = P \left( Z > Z_{\alpha} - \frac{\delta \sqrt{n}}{\sigma} \right)$$

والتي نستنتج منها أن :

$$- Z_{\beta} = Z_{\alpha} - \frac{\delta \sqrt{n}}{\sigma}$$

ولذلك فإن :

$$n = \frac{(Z_{\alpha} + Z_{\beta})^2 \sigma^2}{\delta^2}$$

### ملاحظة هامة

هذه النتيجة صحيحة في الحالة التي يكون فيها الفرض البديل  $\mu < \mu_0$  . في حالة اختبار ثنائي الذيل نحصل على قوة هذا الاختبار  $1 - \beta$  من أجل فرض بديل محدد عندما يكون :

$$n \simeq \frac{(Z_{\alpha/2} + Z_{\beta})^2 \sigma^2}{\delta^2}$$

### مثال (٧,١١)

لاختبار فرض العدم  $\mu = 68$  :  $\mu_0$  ضد الفرض البديل  $\mu > 68$  :  $H_0$  بمستوى من المعنوية قدره  $\alpha = 0.05$  ، وذلك بفرض أن  $\sigma = 5$  ، فإننا نختار عينة من المجتمع المدروس

ونقوم بدراستها أوجد حجم العينة المطلوب إذا كانت قوة هذا الاختبار  $1 - \beta = 0.95$  والوسط الحقيقي هو 69 .

### الحل

نلاحظ أن  $\alpha = \beta = 0.05$  ، وأن  $Z_\alpha = Z_\beta = 1.645$  . ومن أجل الفرض البديل  $\mu = 69$  نأخذ  $\delta = 1$  لذلك فإن :

$$n = \frac{(1.645 + 1.645)^2 (25)}{1} = 270.6$$

هذا يعنى أن الاختبار يتطلب 271 ملاحظة ، وذلك لرفض فرض العدم في 95% من الحالات عندما تكون  $\mu$  أكبر بكثير من 69 . يمكن استخدام نفس الإجراءات لتحديد حجم العينة  $n = n_1 = n_2$  المطلوبة لدى القيام باختبار وسطي مجتمعين بقوة اختبار محددة . فمثلاً لنفرض أننا نرغب في اختبار فرض العدم  $H_0: \mu_1: \mu_2$  ضد الفرض البديل  $H_1: \mu_1: \mu_2$  عندما يكون  $\sigma_1, \sigma_2$  معلومين ، فمن أجل الفرض البديل المحدد  $\mu_1 - \mu_2 = \delta$  تكون قوة اختبارنا محددة بالعلاقة التالية :

$$1 - \beta = P \left[ \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{(\sigma_1^2 + \sigma_2^2)/n}} > Z_{\alpha/2} \mid \mu_1 - \mu_2 = \delta \right]$$

لذلك فإن :

$$\beta = P \left[ - Z_{\alpha/2} < \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{(\sigma_1^2 + \sigma_2^2)/n}} < Z_{\alpha/2} \mid \mu_1 - \mu_2 = \delta \right]$$

$$\beta = P \left[ Z_{\alpha/2} - \frac{\delta}{\sqrt{(\sigma_1^2 + \sigma_2^2)/n}} < \frac{\bar{X}_1 - \bar{X}_2 - \delta}{\sqrt{(\sigma_1^2 + \sigma_2^2)/n}} < Z_{\alpha/2} \mid \mu_1 - \mu_2 = \delta \right]$$

وبالنسبة لهذا الفرض البديل  $\mu_1 - \mu_2 = \delta$  ، فإن للإحصاء  $\frac{\bar{X}_1 - \bar{X}_2 - \delta}{\sqrt{(\sigma_1^2 + \sigma_2^2)/n}}$  توزيعاً طبيعياً معيارياً لذلك فإن :

$$\beta = P \left[ -Z_{\alpha/2} - \frac{\delta}{\sqrt{(\sigma_1^2 + \sigma_2^2)/n}} < Z < Z_{\alpha/2} - \frac{\delta}{\sqrt{(\sigma_1^2 + \sigma_2^2)/n}} \right]$$

من هذه العلاقة الأخيرة نستنتج أن :

$$-Z_{\beta} \approx Z_{\alpha/2} - \frac{\delta}{\sqrt{(\sigma_1^2 + \sigma_2^2)/n}}$$

لذلك فإن :

$$n \approx \frac{(Z_{\alpha/2} + Z_{\beta})^2 (\sigma_1^2 + \sigma_2^2)}{\delta^2}$$

وعبارة حجم العينة المطلوبة من أجل اختبار وحيد الذيل عندما يكون  $n = n_1 = n_2$  تعطى بالعلاقة :

$$n = \frac{(Z_{\alpha/2} + Z_{\beta})^2 (\sigma_1^2 + \sigma_2^2)}{\delta^2}$$

## (٧,٦) الاختبارات المتعلقة بالنسب

### Tests concerning proportions

يتطلب في كثير من مجالات الحياة إجراء اختبارات فروض تتعلق بالنسب ، فمثلا يهتم رجل السياسة عادة في معرفة نسبة الأصوات المؤيدة له في الانتخابات القادمة . كما تهتم جميع الشركات الصناعية بتحديد نسبة العناصر المعيبة وذلك عند شحن كمية من البضائع المصنعة . سندرس مسألة اختبار فرض العدم  $H_0: p = p_0$  وذلك ضد الفرض البديل وحيد الذيل  $H_1: p > p_0$  أو  $H_1: p < p_0$  أو ثنائيا في الذيل  $H_1: p \neq p_0$  وذلك بفرض أن  $p$  يمثل متغير ( وسيط ) التوزيع الحدائي .

إن الإحصاء المناسب والذي سنبنى عليه قرارنا في رفض أو قبول فرض العدم  $H_0$  هو المتغير العشوائي  $X$  . كما يمكن أن نأخذ الإحصاء  $\hat{p} = \frac{X}{n}$  . وستقودنا قيم المتغير  $X$

والتي تكون بعيدة عن الوسط  $\mu = np_0$  إلى رفض الفرض  $H_0$ . ولاختبار الفرض  $H_0: p = p_0$  ضد الفرض البديل  $H_1: p < p_0$  فإننا نستخدم التوزيع الحداني بالتغير  $p = p_0$  و  $q = 1 - p_0$  لتحديد  $P(X \leq x | H_0)$  والقيمة  $x$  تمثل عد النجاحات في عينتان والتي يساوى حجمها  $n$ . إذا كان  $P(X \leq x | H_0) < \alpha$  فيكون اختبارنا على مستوى من المعنوية قدره  $\alpha$  ونرفض عندئذ  $H_0$  لصالح  $H_1$ . وبصورة مشابهة لاختبار الفرض  $H_0: p = p_0$  ضد الفرض البديل  $H_1: p > p_0$ ، فإننا نبحث عن  $P(X \geq x | H_0)$  لصالح  $H_1$  إذا كان هذا الاحتمال أقل من  $\alpha$ . وأخيرا لاختبار الفرض  $H_0: p = p_0$  ضد الفرض البديل  $H_1: p \neq p_0$  بمستوى المعنوية  $\alpha$ ، فإننا نرفض  $H_0$  عندما يكون  $P(X \leq x | H_0) < \alpha/2$  أو إذا كان  $x > np_0$  و  $P(X \geq x | H_0) < \frac{1}{2}$ . يمكن أن نجمل الخطوات المتبعة في اختبار فرض حول النسبة ضد مختلف الفروض البديلة بما يلي:

(١) نحدد فرض العدم  $H_0: p = p_0$

(٢) نحدد الفرض البديل  $H_1: p < p_0$  أو  $p > p_0$  أو  $p \neq p_0$

(٣) نختار مستوى معنوية الاختبار  $\alpha$ .

(٤) نحدد المنطقة الحرجة.

أ — فمن أجل الفرض البديل  $p < p_0$  تكون المنطقة الحرجة ممثلة لجميع قيم  $x$  المحققة للمتبينة  $P(X \leq x | H_0) < \alpha$

ب — أما من أجل  $H_1: p > p_0$  فتؤلف مجموعة القيم  $x$  المحققة للمتبينة  $P(X \geq x | H_0) < \alpha$  صحيح  $x$  ما يسمى بالمنطقة الحرجة في مثل هذه الحالة.

ج — وأخيراً فإن جميع قيم  $x$  المحققة للمتبينة  $P(X \leq x | H_0) < \alpha/2$  وذلك

من أجل  $x < np_0$  أو جميع قيم  $x$  المحققة لـ  $P(X \geq x | H_0) < \alpha/2$  من أجل

$x > np_0$  تؤلف منطقة حرجة للفرض البديل  $H_1: p \neq p_0$

(٥) نجد  $x$  ثم بحسب الاحتمال الموافق

(٦) نتخذ قرار حول رفض  $H_0$  فيما إذا وقع  $x$  في المنطقة الحرجة ونقبله فيما عدا ذلك.

مثال (٧، ١٢)

يدعى صياد أنه يصيب 90% من الطيور التي يطلق عليها عياراته النارية. هل توافق الصياد إدعائه هذا إذا علمت أنه أصاب في يوم من الأيام اثني عشر طيراً عندما

أطلق عليها عشرون طلقة متتالية ؟ استخدم مستوى معنوية قدره  $\alpha = 0.05$  .

### الحل

لحل هذا المثال نتبع الخطوات الست التالية :

(١) نلاحظ أن فرض العدم هي  $H_0: p = 0.9$

(٢) وأن الفرض البديل هي  $H_1: p \neq 0.9$

(٣) أما بالنسبة لمستوى المعنوية فهو  $\alpha = 0.05$

(٤) نلاحظ أن جميع قيم  $x$  المحققة للمتبينة  $0.025 < P(X \leq x | \text{صحيح } H_0)$  تحدد المنطقة الحرجة .

(٥) كما أن لدينا  $x = 12$  وأن  $n = 20$  . لذلك وباستخدام الجدول II نجد أن :

$$P(X \leq 12 | p = 0.9) = \sum_{x=0}^{12} b(x; 20, 0.9) \\ = 0.0004 < 0.025$$

(٦) نرفض الفرض  $H_0$  ونستنتج أن ادعاء هذا الصياد خاطيء .

وجدنا في الفقرة (٤، ٣) أن الاحتمال الحداني قد استنتج صيغة التوزيع الحداني الفعلي أو بواسطة الجدول II ، وذلك في الحالة التي تكون فيها  $n$  صغيرة . أما إذا كانت  $n$  كبيرة فإننا نجرى عمليات التقريب المطلوبة . أما عندما تكون القيمة المختبرة  $p_0$  قريبة من الصفر أو الواحد ، فإننا نستخدم التوزيع اليواسوني بالوسط  $\mu = np$  . هذا ويقدم المنحنى الطبيعي تقريباً يكون عادة مفضلاً في الحالة التي تكون فيها  $n$  كبيرة و  $p_0$  ليست قريبة من الصفر ولا الواحد . باستخدام التقريب الطبيعي فإننا نبني قرارنا على المتغير الطبيعي المعياري :

$$Z = \frac{\hat{p} - p_0}{\sqrt{p_0 q_0 / n}} = \frac{X - np_0}{\sqrt{np_0 q_0}}$$

لذلك فإن المنطقة الحرجة للاختبار ثنائي الذيل بمستوى المعنوية  $\alpha$  هي المنطقة  $Z > Z_{\alpha/2}$  ،



أما بالنسبة للاختبار وحيد الذيل من أجل الفرض البديل  $p < p_0$  فتكون المنطقة  $Z < -Z_{\alpha}$  ممثلة للمنطقة الحرجة في هذا الاختبار . وأخيراً فالنسبة للفرض البديل  $p > p_0$  تكون المنطقة الحرجة هي  $Z > Z_{\alpha}$  .

لاختبار فرض يتعلق بالنسبة وذلك باستخدام منحني التقريب الطبيعي ، فإننا نتبع الخطوات التالية :

(١) فرض العدم  $H_0 : p = p_0$

(٢) الفرض البديل  $p > p_0$  أو  $p < p_0$  أو  $p \neq p_0$   $H_1 :$

(٣) نختار مستوى المعنوية الاختبار  $\alpha$  .

(٤) نحدد المنطقة الحرجة .

أ — إذا كانت الفرض البديل  $p < p_0$   $Z < -Z_{\alpha}$

ب — إذا كانت الفرض البديل  $p > p_0$   $Z > Z_{\alpha}$

ج —  $Z < -Z_{\alpha/2}$  وذلك من أجل الفرض البديل  $p \neq p_0$

(٥) نحسب  $x$  من عينة حجمها  $n$  وبعد ذلك نحسب  $Z = \frac{x - np_0}{\sqrt{np_0q_0}}$

(٦) نتخذ قراراً في رفض  $H_0$  إذا وقع  $Z$  في المنطقة الحرجة ونقبله فيما عدا ذلك .

### مثال (٧، ١٣)

ادعت شركة صناعية أن 95% من عناصرها المنتجة هي جيدة الصنع . وقد قامت الشركة مؤخراً باستخدام وسائل جديدة ومتطورة بقصد خفض نسبة الخسرة بالمائة من العناصر المعيبة في إنتاجها . ثم أنتجت في يوم ما مائة عنصر ، وذلك باستخدام هذه الطرق الحديثة . بعد ذلك قامت بفحصها فوجدت أن ثلاثة عناصر من بينها معيبة . فهل تكون هذه العينة دليلاً كافياً لاستنتاج أن الوسائل الحديثة قد طورت فعلاً عملية الإنتاج هذه ؟ استخدم 0.01 مستوى معنوية .

### الحل

نلاحظ باستخدام الخطوات الست المذكورة سابقاً أن :

(١)  $H_0 : P = 0.95$

(٢)  $H_1 : p > 0.95$

$$\alpha = 0.01 \quad (٣)$$

$$(٤) \text{ المنطقة الحرجة هي } Z > 2.575$$

$$(٥) \text{ لدينا } x = 97 \text{ و } n = 100 \text{ فإن } npq = (100) - (0.95) = 95 \text{ وأن :}$$

$$Z = \frac{97 - 95}{\sqrt{(100)(0.95)(0.05)}} = 0.917$$

(٦) الاستنتاج : بما أن قيمة  $Z = 0.917$  تقع ضمن منطقة القبول ، لذلك فإننا نقبل  $H_0$  ونستنتج أن الطرق الحديثة المستخدمة في عملية الإنتاج لم تطور هذه العملية ولم تخفض نسبة العناصر المعيبة .

### (٧,٧) اختبار الفرق بين نسبتي

#### Testing the difference between two proportions

عندما نرغب في اختبار فرض حول تساوي نسبتي ، فإنه تنشأ عادة حالات متعددة . فمثلاً يمكننا أن نحاول البرهان على أن نسبة أطباء الأطفال في مدينة جدة بالمملكة العربية السعودية يساوي نسبة أطباء الأطفال في مدينة الرياض . كما يمكن لشخص أن يقرر الإقلاع عن التدخين فقط إذا كان متأكداً من أن نسبة المدخنين المصابين بسرطان الرئة يفوق نسبة غير المدخنين المصابين بنفس المرض .

نرغب بشكل عام في اختبار فرض العدم  $H_0 : p_1 = p_2 = p$  ضد فرض بديل مناسبة  $H_1$  . حيث يمثل الوسطاء  $p_1, p_2$  نسبتي المجتمعين ، ويمثل المتغير العشوائى  $\beta_1 - \beta_2$  الإحصاء الذى سنبنى عليه قرارنا . حيث يمثل  $\beta_1, \beta_2$  نسبتي النجاح المحسوبتين من خلال عينتين عشوائيتين مستقلتين حجمهما  $n_1, n_2$  على الترتيب مأخوذتين من مجتمعين حدانيين . وكما مر معنا في (٦,٦) فإن للإحصاء :

$$Z = \frac{\beta_1 - \beta_2}{\sqrt{(p_1q_1/n_1) + (p_2q_2/n_2)}} = \frac{\beta_1 - \beta_2}{\sqrt{pq [(1/n_1) + (1/n_2)]}}$$

توزعاً طبيعياً معيارياً عندما يكون  $H_0$  صحيحاً وحجما العينتين  $n_1, n_2$  كبيرين .

لحساب قيمة  $Z$  علينا أن نقدر المتغير  $p$  الموجود تحت إشارة الجذر . بفرض أن  $x_1, x_2$  يمثلان على الترتيب عدد النجاحات في كل عينة . لذلك يمكن أن نأخذ :

$$\hat{p} = \frac{x_1 + x_2}{n_1 + n_2}$$

وبتبدل  $\hat{p}$  عن  $p$  في عبارة الإحصاء  $Z$  فإننا نجد العلاقة :

$$Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}\hat{q} \{(1/n_1) + (1/n_2)\}}}$$

حيث يمثل  $\hat{q} = 1 - \hat{p}$  .

لذلك وعند اختبار فرض حول تساوى نسبتي مجتمعين عندما تكون أحجام العينات المسحوبة كبيرة فإننا نتبع الخطوات الست التالية :

- (١) نقوم بتحديد فرض العدم  $H_0: p_1 = p_2$
- (٢) بعد ذلك نحدد الفرض البديل  $p_1 > p_2$  أو  $p_1 < p_2$
- (٣) نختار مستوى معنوية الاختبار  $\alpha$
- (٤) فتكون المنطقة الحرجة هي إحدى المنطقتين التاليتين :
  - أ —  $Z < -Z_{\alpha}$  وذلك من أجل الفرض البديل  $p_1 < p_2$
  - ب —  $Z > Z_{\alpha}$  وذلك من أجل الفرض البديل  $p_1 > p_2$
  - ج —  $Z < -Z_{\alpha/2}$  و  $Z > Z_{\alpha/2}$  وذلك إذا كان الفرض البديل  $p_1 \neq p_2$

(٥) ثم نحسب  $\hat{p}_1 = \frac{x_1}{n_1}$  ,  $\hat{p}_2 = \frac{x_2}{n_2}$  و  $\hat{p} = \frac{x_1 + x_2}{n_1 + n_2}$  وبعدها نجد قيمة الإحصاء :

$$Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}\hat{q} \{(1/n_1) + (1/n_2)\}}}$$

- (٦) وأخيراً فإننا نتخذ قرارنا حول رفض  $H_0$  فيما لو وقعت قيمة الإحصاء  $Z$  المحسوبة سابقاً داخل المنطقة الحرجة أو أن نقبله في غير ذلك .

## مثال (٧, ١٤)

يجرى التصويت على اقتراح لإقامة مصنع كيميائي بين المقيمين في مدينة معينة والمقاطعات المحيطة بها . وسيكون المكان الذى سينشأ عليه هذا المصنع ضمن مشروع الاقتراح هو فى داخل حدود المدينة . ولهذا السبب فإن المصوتين من بين سكان المقاطعات يشعرون أن الاقتراح سيمر بسبب النسبة الكبيرة لأصوات المقترعين فى المدينة والذين يدعمون مثل هذا المشروع .

ولتحديد ما إذا كان هناك فرق هام بين نسبة أصوات المقيمين فى المدينة وأصوات المقيمين فى المقاطعات والذين يدعمون الاقتراح ، فقد جرى اقتراح جزئى ، فتبين أن 140 من أصل 250 من المقترعين فى المدينة قد صوتوا لصالح هذا الاقتراح ، وأن 290 من 600 مقترح فى المقاطعات قد صوتوا لصالحه أيضا . فهل نوافق القول بأن نسبة أصوات المقيمين فى المدينة والمؤيدين للاقتراح المذكور هى أكبر من نسبة أصوات أهل المقاطعات ؟ استخدم 0.025 مستوى معنوية .

## الحل

نفرض أن  $p_1$  ,  $p_2$  يمثلان على الترتيب نسبة المقترعين فى المدينة والمقاطعات والمصوتين لصالح الاقتراح . نلاحظ باتباع الخطوات الست المذكورة سابقا أن :

$$H_0 : p_1 = p_2 \quad (١)$$

$$H_1 : p_1 > p_2 \quad (٢)$$

$$\alpha = 0.025 \quad (٣)$$

$$(٤) \text{ أما بالنسبة للمنطقة الحرجة فهي } Z > 1.96$$

$$(٥) \text{ نحسب من معطيات المسألة } \hat{p}_1 = \frac{x_1}{n_1} = \frac{140}{250} = 0.56 , \hat{p}_2 = \frac{x_2}{n_2} = \frac{290}{600} = 0.483$$

$$\bar{p} = \frac{x_1 + x_2}{n_1 + n_2} = 0.5058$$

لذلك فإن :

$$Z = \frac{0.56 - 0.483}{\sqrt{(0.5058)(0.4942) \left[ \frac{1}{250} + \frac{1}{600} \right]}} = 2.045$$

(٦) نلاحظ أن قيمة الإحصاء  $Z = 2.045$  تقع داخل المنطقة الحرجة ، ولذلك فإننا نرفض الفرض  $H_0$  ونوافق القول بأن نسبة أصوات المقترعين فى المدينة لصالح الاقتراح أكبر من نسبة أصوات أهل المقاطعات المحيطة بهذه المدينة .

## تمارين محلولة

## تمرين (١)

من المعلوم أن مقياس الذكاء في إحدى مدارس بحرة التابعة لمدينة جدة في المملكة العربية السعودية يتوزع وفقاً للتوزيع الطبيعي بالوسط  $\mu = 110$  والتباين  $\sigma^2 = 100$ . ولدراسة مدى اختلاف وسط الذكاء اخترنا عينة من إحدى مدارس جدة مؤلفة من 25 طالباً. فتبين أن وسط الذكاء  $\bar{x} = 115$ . فهل وسط الذكاء في هذه المدرسة يختلف عن وسط الذكاء في بحرة؟ استخدم لذلك مستوى معنوية  $\alpha = 0.1$ .

## الحل

لحل هذا التمرين نتبع الخطوات الست التالية :

(١) لنختبر فرض العدم  $H_0: \mu \neq 110$

(٢) ضد الفرض البديل  $H_1: \mu_x \neq 100$

(٣) وذلك تحت مستوى من المعنوية قدره  $\alpha = 0.1$

(٤) نلاحظ بالعودة إلى السطر الأول في الجدول (٧،٢) أن المنطقة الحرجة للفرضية  $H_0$

تحدد بالمتباينين  $Z > Z_{\alpha/2}$  و  $Z < -Z_{\alpha/2}$ . من الجدول IV نجد أن هذه المنطقة هي

$$Z < -1.65, Z > 1.65$$

(٥) لنحدد قيمة احصاء الاختبار :

$$Z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} = \frac{115 - 110}{10 / \sqrt{25}} = 2.5$$

(٦) الاستنتاج : بما أن هذه القيمة تقع ضمن المنطقة الحرجة ، لذلك فإننا نرفض

الفرض  $H_0$  ونستنتج أن وسط الذكاء في هذه المدرسة يختلف عنه في منطقة بحرة .

## تمرين (٢)

بفرض أن وسط الزمن اللازم للقيام بعملية صناعية معينة هو  $\mu = 12.5$  دقيقة .  
اخترنا عشرة مستخدمين جدد ودربوا على هذه العملية ، ولدى اختبار الزمن اللازم  
لكل منهم للقيام بمثل هذه العملية وجدنا أن :

رقم العامل	1	2	3	4	5	6	7	8	9	10
الزمن اللازم	9.3	12.1	15.7	10.3	12.2	14.8	15.1	13.2	15.9	14.5

اختبر الفرض التالي : أن وسط الزمن اللازم لأى مستخدم لا يختلف عن وسط الزمن  
12.5 دقيقة . استخدم 0.05 مستوى معنوية .

## الحل

لحل هذا التمرين نتبع الخطوات الست التالية :

(١) أن فرض العدم فهو  $H_0: \mu = 12.5 (= \mu_0)$

(٢) أما بالنسبة للفرض البديل فهو  $H_1: \mu \neq 12.5$

(٣) أن مستوى المعنوية  $\alpha = 0.05$  .

(٤) لنحدد المنطقة الحرجة لإحصاء الاختبار  $t = \frac{\bar{x} - \mu}{S / \sqrt{n}}$  وذلك من أجل عدد من

درجات الحرية قدره  $9 = n - 1$  . وبالعودة إلى الأسطر الثاني من الجدول (٧،٢) نجد أن  
المنطقة الحرجة للفرض  $H_0$  هي المنطقة .

$$t < -t_{1-\alpha/2} = t_{0.975} = -2.26 , t > t_{\alpha/2} = t_{0.025} = 2.262$$

(٥) نلاحظ من معطيات المسألة أن  $\bar{x} = 13.31$  وأن  $S = 2.28$  . لذلك فإن قيمة  
إحصاء الاختبار هي :

$$T = \frac{\bar{x} - \mu}{S / \sqrt{n}} = \frac{13.31 - 12.5}{(2.28) / \sqrt{10}} = 1.13$$

(٦) بما أن قيمة  $T = 1.13$  لا تقع ضمن المنطقة الحرجة ، لذلك فإننا نقبل  $H_0$  ونستنتج أنه لا يوجد دليل على أن وسط الزمن اللازم للعمال الجدد للقيام بمثل هذه العملية يختلف عن الوسط المفروض ( 12.5 دقيقة ) .

### تمرين (٣)

يدعى صاحب مصنع لإنتاج الحبال أن لحباله المنتجة وسط مقاومة للقطع قدره 8000 N ، ولدراسة هذا الادعاء قمنا باختيار ستة حبال من إنتاج هذا المصنع ، فظهرت نتائج التجربة أن لها وسط قوة مقاومة للقطع قدره 7750 N وانحرافا معياريا 145 N ، فهل يمكن تأييد ادعاء صاحب المصنع ؟ استخدم مستوى معنوية قدره 0.05 .

### الحل

علينا أن نقرر بين الفرضين :

$H_0: \mu = 8000$  N وادعاء المصنع له ما يبرره

$H_1: \mu < 8000$  N وادعاء المصنع ليس له ما يبرزه

أى أن المطلوب اختبار فرض وحيد الذيل . ونحن نجد تحت الفرض  $H_0$  أن :

$$T = \frac{\bar{X} - \mu}{S / \sqrt{n}} = \frac{7750 - 8000}{145 / \sqrt{6}} = - 4.22$$

وعند مستوى من المعنوية قدره  $\alpha = 0.05$  ، أن المنطقة الحرجة للفرض  $H_0$  حسب الجدول (٧، ٢) السطر الثانى هى المنطقة  $t_{0.05} = - 2.01$  ، وذلك من أجل  $n - 1 = 5$  . وبما أن قيمة الإحصاء  $t = - 4.22$  تقع داخل المنطقة الحرجة ، لذلك فإننا نرفض  $H_0$  وهذا يعين أن إدعاء صاحب المصنع ليس له ما يبرره .

### تمرين (٤)

قمنا بتجربة لتقييم طريقة جديدة فى إنتاج الماس . وقد ولدنا ست ماسات بالطريقة الجديدة فكانت أوزانها 0.57, 0.54, 0.45, 0.61, 0.52, 0.48 . وقد بينت دراسة كلفة هذه الطريقة أنه يجب أن يكون وسط وزن الماس المستحصل أكبر من 0.5

قيراطاً لكي تكون الطريقة الجديدة مربحة . هل تقدم الأوزان السابقة دلالة كافية للقول بأن الطريقة الجديدة مقبولة ؟ استخدم مستوى معنوية  $\alpha = 0.05$  .

### الحل

حل هذا التمرين نقوم باتباع الخطوات الست المعروفة :

- (١) أن فرض العدم هو  $H_0: \mu = 0.5$
- (٢) كما أن الفرض البديل هو  $H_1: \mu > 0.5$
- (٣) ومستوى معنوية الاختبار  $\alpha = 0.05$
- (٤) وحسب السطر الثاني في الجدول (٧,٢) فإننا نجد أن المنطقة الحرجة هي المنطقة  $t_{0.05} > 2.015$  ، وحيث أن  $n = 5$  . وبالعودة إلى الجدول ٧ نجد أن هذه المنطقة هي  $t < 2.015$
- (٥) ونلاحظ أن قيمة إحصاء الاختبار هي :

$$t = \frac{\bar{X} - \mu_0}{S / \sqrt{n}} = \frac{0.53 - 0.5}{0.0559 / \sqrt{6}} = 1.31$$

- (٦) الاستنتاج : نلاحظ أن قيمة الإحصاء المحسوب  $t = 1.31$  تقع خارج منطقة الرفض ، لذلك فإننا نقبل الفرض  $H_0: \mu = 0.5$  ، وهذا يعني أن المعلومات التي حصلنا عليها لا تقدم لنا دليلاً كافياً على أن وسط وزن الماس المنتج يتجاوز 0.5 قيراطاً .

### تمرين (٥)

تحتاج عملية تجميع جهاز تلفزيون في منشأة النصر لصنع التلفزيونات إلى تدريب لمدة شهر تقريباً يتلقاه مستخدم جديد للوصول إلى كفاءة عالية . وقد اقترحت طريقة جديدة للتدريب . وتم القيام بتجربة لمقارنة الطريقة الجديدة بالطريقة المعتادة . فقمنا بتدريب مجموعتين من الشباب ، تحوى كل منهما تسعة عمال جدد مستخدمين ، واحدة منهما في كل من الطريقتين ، وذلك لمدة ثلاثة أسابيع . ثم قسنا في النهاية الزمن اللازم بالدقائق لكل عامل لتجميع الجهاز ، فوجدنا النتائج المسجلة في الجدول التالي :



الطريقة القديمة	الطريقة الجديدة
35	32
31	37
29	35
25	28
34	41
40	44
27	35
32	31
31	34

هل تقدم هذه النتائج معلومات كافية على تفوق الطريقة الجديدة ؟

**الحل**

لنفرض أن الزمن اللازم لتجميع جهاز تلفزيوني بالطريقة القديمة يتوزع وفقا للتوزيع الطبيعي بالوسط  $\mu_1$  والتباين  $\sigma^2$  وبالنسبة للطريقة الجديدة بالوسط  $\mu_2$  والتباين  $\sigma^2$ .

من الجدول السابق نجد أن  $\bar{x}_1 = 35.22$  ،  $\bar{x}_2 = 31.56$

$$\sum_{j=1}^9 (\bar{x}_j - \bar{x}_1)^2 = 195.56 , \quad \sum_{i=1}^9 (\bar{x}_i - \bar{x}_2)^2 = 16.22$$

أما تقدير التباين المشترك فهو يساوى :

$$S^2 = \frac{\sum_{j=1}^9 (x_j - \bar{x}_1)^2 + \sum_{i=1}^9 (\bar{x}_i - \bar{x}_2)^2}{n_1 + n_2 - 2} = \frac{195.56 + 160.22}{9 + 9 - 2} = 22.24$$

وقيمة S هي 4.71

لنتخبر فرض العدم  $H_0: \mu_1 - \mu_2 = 0$  ضد الفرض البديل  $H_1: \mu_1 - \mu_2 > 0$  . من السطر الرابع في الجدول (٧، ١) وبالعودة إلى الجدول ٧ نجد أن القيمة الحرجة لـ  $t$  من أجل  $\alpha = 0.05$  ،  $\nu = n_1 + n_2 - 2 = 16$  ، درجة حرية هي القيمة  $t = 1.746$  ، وتكون المنطقة الحرجة  $t > 1.746$  . وإذا حسبنا قيمة إحصاء الاختبار لوجدنا أن :

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{35.22 - 31.56}{4.71 \sqrt{\frac{1}{9} + \frac{1}{9}}} = 1.65$$

نلاحظ أن قيمة هذا الإحصاء  $t = 1.65$  لا تقع ضمن المنطقة الحرجة ، ولذلك فإننا نقبل  $H_0$  ونستنتج أنه لا توجد دلالة كافية على أن الطريقة الجديدة متفوقة على القديمة .

### تمرين (٦)

يريد مكتب الدراسات في جامعة الملك عبد العزيز بمجدة دراسة ما إذا كانت الحالة الاقتصادية للطلاب ذات أثر في قدراته الدراسية . ول هذه الغاية تم تقسيم الطلاب إلى فئتين وأخذت عينة عشوائية مؤلفة من مائة طالب من كل فئة وتم دراستها ، فكان وسط العينة الأولى  $\bar{x} = 67.5$  وتباينها العيني  $S_1^2 = 225$  . كما وجد أن  $\bar{x}_2 = 63.5$  وأن  $S_2^2 = 250$  . هل تقدم هذه المعلومات دلالة كافية على وجود فرق في القدرات الدراسية بين الفئتين ؟ استخدم لذلك مستوى معنوية  $\alpha = 0.05$  .

### الحل

نرمز للمعدل العام لطلاب الفئة الأولى بالرمز  $\mu_1$  ، ولطلاب الفئة الثانية بالرمز  $\mu_2$  . لتتبع الخطوات الستة المعروفة :

(١) أن فرض العدم هو  $H_0 : \mu_1 - \mu_2 = d_0$  ( $d_0 = 0$ )

(٢) أما الفرض البديل فهو  $H_1 : \mu_1 - \mu_2 \neq d_0$

(٣) بالنسبة لمستوى معنوية الاختبار فهو يساوي  $\alpha = 0.05$

(٤) نعلم أن الإحصاء  $\bar{x}_1 - \bar{x}_2$  يمثل تقديراً جيداً للفرق  $\mu_1 - \mu_2$  وله توزيع قريب من التوزيع الطبيعي ( لأن  $n_1 = 100, n_2 = 100$  ) بوسط قدره  $\mu_1 - \mu_2$  وتباين قدره :

$$\sigma_{\bar{x}_1 - \bar{x}_2}^2 = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$$

بالعودة إلى السطر الثالث في الجدول (٧،٢) نجد أن المنطقة الحرجة للإحصاء :

$$Z = \frac{\bar{X}_1 - \bar{X}_2 - d_0}{\sqrt{(\sigma_1^2/n_1) + (\sigma_2^2/n_2)}} \quad \text{حيث المنطقة } Z < -1.96, Z > 1.96.$$

(٥) لنحسب قيمة إحصاء الاختبار فنجد أن :

$$Z = \frac{67.5 - 63.5}{\sqrt{(225/100) + (250/100)}} = 1.83$$

(٦) بما أن قيمة هذا الإحصاء لا تقع ضمن المنطقة الحرجة ، لذلك فإننا نقبل  $H_0$  ونستنتج أن الحالة الاقتصادية لا علاقة لها بقدرات الطالب الدراسية تحت مستوى معنوية قدره  $\alpha = 0.05$ .

تمرين (٧)

تبين سجلات مستشفى جامعة الملك عبد العزيز بمكة أن 52 رجلاً من عينة مكونة من 1000 رجل يقابلها 23 امرأة من أصل 1000 امرأة ممن كانوا يعانون من مرض في المعدة خلال العام ١٤٠٠ - ١٤٠١ هجرية . هل تقدم هذه المعلومات الإحصائية دلالة كافية على أن نسبة المرضى بين الرجال بأمراض المعدة هو أكبر منها عند النساء ؟

الحل

نفرض أن عدد المرضى الوافدين إلى قسم الأمراض المعدية في المستشفى خلال عام سواء من الرجال أو النساء يتبع التوزيع الحداني بوسطين هما على الترتيب  $p_1, p_2$  لاختبار فرض العدم  $H_0 : p_1 = p_2$  ضد الفرض البديل  $H_1 : p_1 \neq p_2$  وكتقدير للفرض  $p_1 - p_2$  نأخذ التقدير النقلي  $\hat{p}_1 - \hat{p}_2$  . نعلم أن لهذا التقدير توزيعاً طبيعياً بالوسط  $p_1 - p_2$  والتباين  $\sigma_{(\hat{p}_1 - \hat{p}_2)}^2 = \frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}$  و  $H_0$  نأخذ المتغير العشوائي

$$Z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sigma_{(\hat{p}_1 - \hat{p}_2)}}$$

فإذا استخدمنا قيم تقريبية مناسبة لكل من  $p_1, p_2$  في عبارة  $\sigma_{(\hat{p}_1 - \hat{p}_2)}^2$  فإننا نجد أن التقدير المشترك هو :

$$\bar{p} = \frac{x_1 + x_2}{n_1 + n_2} = \frac{52 + 23}{1000 + 1000} = 0.0375$$

وإحصاء الاختبار هو :

$$Z = \frac{\bar{p}_1 - \bar{p}_2}{\sqrt{\hat{p}\hat{q} \left( \frac{1}{n_1} + \frac{1}{n_2} \right)}} = \frac{0.52 - 0.023}{\sqrt{(0.0375)(0.9625) \left( \frac{1}{1000} + \frac{1}{1000} \right)}}$$

وتكون منطقة الرفض من أجل  $\alpha = 0.05$  هي المنطقة  $Z > 1.96$  و  $Z < -1.96$ . نلاحظ أن قيمة الإحصاء السابقة تقع ضمن هذه المنطقة الحرجة لذلك نرفض الفرض  $p_1 = p_2$  ونستنتج أن المعلومات الإحصائية السابقة تقدم لنا دلالة قوية على أن نسبة المرض بأمراض المعدة من الرجال كانت أكبر منها عند النساء ممن راجعوا المستشفى المذكور خلال نفس العام .

#### تمرين (٨)

جرى اختبار لدراسة تأثير سماد من نوع معين على إنتاج الشعير . ولهذا الغرض اختيرت 24 قطعة من الأرض لها نفس المساحة ، و عولج نصفها بالسماد بينما ترك النصف الآخر بدون معالجة . ولدى جني المحصول وجد أن وسط الغلة من الشعير في مجموعة القطع المتروكة هو 4.8 طناً ، وأن انحرافها المعياري هو 4 طن ، بينما كان وسط غلة الفدان للقطع التي تمت معالجتها بالسماد هو 5.1 طناً وانحرافها المعياري 3.6 طناً . هل هناك تحسن معنوي في إنتاج الشعير نتيجة استخدام السماد ، وذلك باستخدام مستوى معنوية قدره  $\alpha = 0.01$  ؟

#### الحل

نفرض أن  $\mu_1, \mu_2$  تمثل وسط مجتمع غلة الشعير من الأرض المعالجة والأرض غير المعالجة على الترتيب ، علينا أن نقرر بين الفرضين :

$$H_0 : \mu_1 = \mu_2$$

$$H_1 : \mu_1 > \mu_2$$

والسماد يؤدي إلى تحسين الغلة ونحت فرض العدم  $H_0$  نأخذ الإحصاء

$$T = \frac{\bar{X}_1 - \bar{X}_2}{S_p \cdot \sqrt{1/n_1 + 1/n_2}} \text{ من أجل } 22 = n_1 + n_2 - 2 = \text{درجة من الحرية ، وحيث}$$

$$S_p^2 = \frac{(n_1 - 1) S_1^2 + (n_2 - 1) S_2^2}{n_1 + n_2 - 2} \text{ تمثل}$$

وهكذا نجد أن :

$$S_p^2 = \frac{(11)(4)^2 + (11)(3.6)^2}{12 + 12 - 2} = 14.48$$

أما قيمة إحصاء الاختبار T فهي تساوى :

$$T = \frac{5.1 - 4.8}{(3.80) \sqrt{1/12 + 1/12}} = 0.193$$

من الجدول (٧،٢) السطر الرابع نجد أن المنطقة الحرجة للفرض  $H_0$  هي  $t_{0.01} = 22$  درجة من الحرية ، وهذه المنطقة هي  $T > 2.508$  . وبما أن قيمة الإحصاء  $T = 0.193$  لا تقع ضمن المنطقة الحرجة لذلك فإننا نقبل  $H_0$  .

تمرين (٩)

كان الانحراف المعياري في فترات سابقة لأوزان عبوات وزن 40.0 N تملأ بواسطة آلة معينة هو 0.25N ولدراسة الزيادة الظاهرة في التباين اخترنا عينة عشوائية مؤلفة من 20 عبوة ، فكان انحرافها المياري 0.32N . فهل هذا يدل على زيادة ظاهرة في التباين ؟ استخدم مستوى معنوية قدره 0.05 .

الحل

لحل هذا التمرين نتبع الخطوات الست المعروفة :

(١) أن فرض العدم هو  $H_0 : \sigma = 0.25$

- (٢) أما الفرض البديل فهو  $H_1: \sigma > 0.25$  وهذا يعنى أن هناك زيادة في التباين .  
 (٣) والنسبة لمستوى معنوية الاختبار فهو  $\alpha = 0.05$   
 (٤) وحسب الجدول (٧،٢) السطر السادس نجد أن المنطقة الحرجة في هذا الاختبار هي المنطقة المحددة بالتباينة  $\chi^2 > \chi^2_{\alpha}$  علماً بأن عدد درجات الحرية لتوزيع كاي - مربع هو  $\nu = n - 1 = 19$  وعلى هذا فإن المنطقة الحرجة هي  $\chi^2 > 30.144$ . وذلك بالعودة إلى الجدول VI .  
 (٥) لنحسب قيمة إحصاء الاختبار :

$$\chi^2 = \frac{(n - 1) S^2}{\sigma_0^2} = \frac{(19) (0.32)^2}{(0.25)^2} = 31.2$$

- (٦) الاستنتاج : بما أن قيمة إحصاء الاختبار تقع داخل المنطقة الحرجة ، لذلك فإننا نرفض  $H_0$  بمستوى معنوية  $\alpha = 0.05$  ونقبل  $H_1$  ، وهذا يعنى أن هناك زيادة ظاهرة في التباين .

### تمرين (١٠)

عيتان عشوائيتان مستقلتان حجماهما  $n_2 = 10, n_1 = 8$  مأخوذتان من مجتمعين تباينهما  $\sigma_1^2, \sigma_2^2$  على الترتيب . حسبنا من خلال هاتين العينتين كلا من  $S_1^2, S_2^2$  فوجدنا أن  $S_1^2 = 7.14, S_2^2 = 3.21$  . هل تقدم هذه القيم دلالة كافية على عدم تساوى تباينى المجتمعين المذكورين ؟ افرض أن توزع المجتمعين المدروسين طبيعيين .

### الحل

لحل هذا التمرين نتبع الخطوات الست المعروفة لدراسة اختبار ما

- (١) نفرض أن  $\sigma_1^2 = \sigma_2^2$   $H_0$   
 (٢) وأن الفرض البديل هو  $H_1: \sigma_1^2 \neq \sigma_2^2$   
 (٣) لنفرض أن مستوى معنوية الاختبار هو  $\alpha = 0.1$   
 (٤) بالعودة إلى الجدول (٧،٢) السطر الأخير نجد أن المنطقة  $(\nu_1, \nu_2)$   $F < f_{\alpha/2}$  و  $F > f_{\alpha/2}$  تحدد لنا ما يسمى بالمنطقة الحرجة حيث رمزنا لإحصاء الاختبار بالرمز  $F = \frac{S_1^2}{S_2^2}$  حيث يمثل  $\nu_1 = n_1 - 1 = 7$  و  $\nu_2 = n_2 - 1 = 9$  ، وبالعودة إلى الجدول VII نجد أن

$$f_{0.05}(v_1, v_2) = f_{0.05}(7, 9) = 3.29$$

وبما أن  $f_{1-\alpha}(v_1, v_2) = \frac{1}{f_{\alpha}(v_2, v_1)}$  فمن الجدول VII نجد أيضا أن :

$$f_{1-\alpha/2}(v_1, v_2) = f_{0.95}(7, 9) = \frac{1}{f_{0.05}(9, 7)} = \frac{1}{3.68} = 0.27$$

والمنطقة الحرجة في هذا الاختبار هي  $F > 3.29$  ,  $F < 0.27$  نلاحظ أن قيمة إحصاء الاختبار :

$$F = \frac{S_1^2}{S_2^2} = \frac{7.14}{3.21} = 2.22$$

وحيث أن هذه القيمة لا تقع ضمن منطقة الرفض ، لذلك نقبل فرض العدم ، وهذا يعني أنه لا يوجد دلالة كافية على وجود فرق بين التباينين .





# الفصل الثامن

## الانحدار والاتّباط

- تمهيد
- الانحدار الخطي
- الانحدار الخطي البسيط
- خواص
- تقديرات المربعات الصغرى
- نهايات الثقة واعتبارات المنووية
- تحليل
- القياسات المتكررة
- لا
- الارتباط
- التباين



## (٨,١) تمهيد

سندرس في هذا الفصل السلوك المشترك لمتغيرين أو أكثر ، فغالباً ما يهتم باحث ما أو مجرب بدراسة ما إذا كانت هناك علاقة بين متغيرين أو أكثر ، ومن ثم يلجأ إلى التعبير عن هذه العلاقة بصورة رياضية وذلك بتحديد المعادلة التي تربط هذه المتغيرات .

فمثلاً يدرس الفيزيائي العلاقة بين ضغط وزن معين من الغاز ( والذي يعتمد على درجة حرارته ) وحجم الغاز ، وكذلك يدرس المهندس الزراعي العلاقة بين كمية محصول غلة القمح وكمية السماد التي يتم بها تزويد نبات القمح في نفس العام ، كما يهتم إدارات الجامعات بمسألة الإنجاز المتوقع لطالب عند نهاية سنته الجامعية الأولى ، فقد ترغب جامعة ما في تقدير ما سيكون عليه معدل كل طالب في سنته الجامعية الأولى وذلك قبل التسجيل . والملاحظ أن معدل الطالب سيكون تابعاً لعدة متغيرات . مثلاً الدرجة التي حصل عليها هذا الطالب خلال امتحان الثانوية في مادة معينة ، مرتبة نجاحه في الثانوية ، وكذلك درجة الاختبار التمهيدى الذى تمجربه الجامعة في نفس المادة .

نلاحظ أن جميع المسائل المطروحة سابقاً تمثل مسائل طبيعية عامة جداً ، فنحن نهتم بالعلاقة بين متغير عشوائى  $y$  - مثلاً - وعدد من المتغيرات المستقلة وغير العشوائية في طبيعتها ، مثل  $x_1, x_2, \dots, x_n$  . ففى المثال الأخير كان المتغير العشوائى  $y$  ممثلاً لمعدل الطالب في امتحان الرياضيات مثلاً ، أما المتغيرات  $x_1, x_2, x_3$  فهى على الترتيب درجة الرياضيات في امتحان الشهادة الثانوية ، مرتبة الطالب ، ودرجة الطالب في الامتحان التمهيدى . إن هدف الجامعة هو قياس  $x_1, x_2, x_3$  من أجل طالب معين ، ثم تبديل هذه القياسات في معادلة تنبؤ تحاول الحصول عليها لتحصل بذلك على تنبؤ عما سيكون عليه معدل الطالب في نهاية السنة الجامعية الأولى . إذا لا بد قبل كل شئ من تحديد المتغيرات المستقلة  $x_1, \dots, x_n$  . ثم الحصول على قياس لمدى علاقتها بالمتغير  $y$  ، وبعد ذلك وضع معادلة تنبؤ تعبر عن  $y$  كدالة بهذه المتغيرات المستقلة . تُعرف مسألة تنبؤ متغير بمعرفة متغيرات أخرى بمسألة الانحدار .

## (٨,٢) الانحدار الخطي The linear regression

### تعريف (٨,١) البيانات الثنائية

بفرض أن  $x_1, x_2, \dots, x_n$  تمثل قيم متغير  $x$  ،  $y_1, y_2, \dots, y_n$  القيم المقابلة لمتغير ثان  $y$  ، عندئذ تسمى مجموعة الأزواج :  $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$  بيانات ثنائية .

### مثال (٨,١)

لنفرض أن القيم  $x_1, x_2, \dots, x_{25}$  تمثل أوزان 25 شخصا في مدينة جدة . ولنفرض أن  $y_1, y_2, \dots, y_{25}$  تمثل أوزان أكبر أولاد هؤلاء الأشخاص على الترتيب ، عندئذ تمثل مجموعة الأزواج  $(x_1, y_1), \dots, (x_{25}, y_{25})$  بيانات ثنائية .

### مثال (٨,٢)

بفرض أن القيم  $x_1, \dots, x_{10}$  تمثل الدرجة التى نالها كل طالب ( من بين كل عشرة طلاب ) فى مادة الرياضيات فى امتحان الشهادة الثانوية ، وأن  $y_1, y_2, \dots, y_{10}$  تمثل القيم الموافقة التى حصل عليها كل طالب من هؤلاء الطلاب خلال الامتحان التمهيدى الذى أجرته الجامعة فى هذه المادة عند فحص القبول . عندئذ تمثل الأزواج  $(x_1, y_1), \dots, (x_{10}, y_{10})$  مجموعة بيانات ثنائية .

لدراسة العلاقة الممكنة مثلا بين درجة الرياضيات  $x$  فى امتحان الثانوية ، ودرجة الرياضيات فى الفحص التمهيدى فى الجامعة ، فإننا نقوم بدراسة العلاقة الممكنة بين أزواج قيم  $x$  وقيم  $y$  ، وذلك بنشر هذه الأزواج فى أماكنها المحددة فى المستوى الاحداثى . وعندئذ يمثل المخطط الناتج عن هذه النقاط ما يسمى بمخطط الانتشار (scatter diagram) .

إن إحدى الطرق الممكنة للحصول على معادلة تنبؤ تربط بين  $x$  و  $y$  هو وضع مسطرة فوق التمثيل البياني ثم تحريكها حتى تبدو وكأنها تمر عبر أكبر عدد ممكن من النقاط ، وهذا الوضع يقدم لنا ما يسمى بأفضل تلاؤم (best fit) مع المعلومات

الإحصائية المتوافرة . وعند رسم هذا المستقيم عبر هذه النقاط تكون مسألة التنبؤ المطروحة منتهية ، حيث إنه بإمكاننا استخدام هذا الخط للتنبؤ بدرجة الطالب  $y$  في الرياضيات عند انتهاء السنة الأولى ، وذلك من أجل قيمة معينة لدرجته  $x$  في اختبار الثانوية .

نلاحظ في المثال السابق أننا اخترنا نموذجاً رياضياً يعبر عن العلاقة المفروضة بين  $x$  و  $y$  ، ومن المعروف أنه بإمكاننا التعبير عن معادلة أى مستقيم على الشكل :

$$y = a + bx$$

حيث يمثل  $a$  ترتيب نقطة تقاطع هذا المستقيم مع المحور الرأسي ، كما يمثل  $b$  ميل هذا المستقيم ، وكما نعلم فإنه يقابل كل مستقيم معادلة خطية بسيطة من هذا النوع والعكس صحيح . والملاحظ أنه عندما نقوم برسم مستقيم عبر الأزواج نكون قد اخترنا وبشكل آلي معادلة رياضية .

$$y = a + bx$$

وبهذا الشكل تحدد قيمتا  $a$  و  $b$  بصورة وحيدة . نسمى النموذج الخطي  $y = a + bx$  بنموذج رياضي حتمى لأنه عندما نبدل قيمة  $x$  في المعادلة نحصل على قيمة محددة لـ  $y$  دون أن يكون هناك أى مجال للخطأ .

إن النماذج الحتمية ملائمة لشرح كثير من الظواهر الفيزيائية والتنبؤ بها عندما يكون خطأ هذا التنبؤ مهماً من وجهة النظر العملية . فمثلاً يسمح لنا قانون نيوتن  $F = m.a$  ( حيث تمثل  $F$  القوة التى ينتجها جسم متحرك كتلته  $m$  ، وتسارعه  $a$  ) فى كثير من الحالات العملية بالتنبؤ بالقوة  $F$  بخطأ صغير مهمل عملياً ، واعتبار الخطأ صغيراً أو كبيراً هو مسألة نسبية . فقد يكون خطأ قدره نيوتن واحداً صغيراً جداً عند إنشاء أساس الجسر ، إلا أنه خطأ كبير جداً بالنسبة لصنع قاعدة إطلاق صاروخ متجه إلى القمر . ولا يمكننا تجاهل هذا الخطأ فى عدد كبير من التجارب الفيزيائية ، وبهذا فإننا سنتردد كثيراً جداً فى منح الكثير من الثقة لتنبؤ غير مصحوب بقياس لجودة هذا التنبؤ . ولهذا فإن طريقة المسطرة والنظر لاختبار خط مستقيم يربط بين قيم  $x$  و  $y$  هى طريقة غير مقبولة ، ومحدودة الفائدة .

لتفسير بعض الظواهر الفيزيائية يمكننا استخدام النماذج الرياضية الاحتمالية عوضاً عن النماذج الحتمية ، ومن المعلوم أن النموذج الرياضى الاحتمالى يحوى متغيراً أو أكثر من المتغيرات ذات الطبيعة العشوائية والتي لها توزيعات احتمالية محددة .

من الملائم تعريف المتغير العشوائى  $Y$  الموافق لقيمة ثابتة لـ  $x$  والذي نرمز له بالرمز  $Y_x$  ، ولتوزيعه الاحتمالى بالرمز  $f(Y_x)$  . من الواضح أنه إذا كان  $x = x_i$  فنعدئذ يمثل الرمز  $Y_{x_i}$  المتغير العشوائى  $Y_i$  . أن تعبير الانحدار الخطى يدل حتماً على أن وسط  $Y_x$  ارتباطاً خطياً بـ  $x$  بواسطة علاقة ميل التقاطع العادية .

$$\mu_{Y_x} = \lambda + \beta x$$

حيث يمثل كلاً من  $\lambda$  و  $\beta$  وسيطين يطلب تقديرهما من بيانات العينة . فإذا رمزنا لتقديرى هذين الوسيطين بالرمزين  $a$  و  $b$  على الترتيب ، فنعدئذ يكون تقدير المتغير المرتبط  $y$  والذي نرمز له بالرمز  $\hat{y}$  محددًا بعلاقة خط الانحدار العينية .

$$\hat{y} = a + bx$$

### (٨,٣) الانحدار الخطى البسيط The simple linear regression

سنحاول فى هذه الفقرة تطوير إجراءات إيجاد تقدير ثوابت الانحدار  $\lambda$  و  $\beta$  ، وذلك من أجل جعل معادلة الانحدار قابلة للاستخدام فى معادلة التنبؤ أو فى معادلة تقدير وسط  $Y$  من أجل قيمة معينة للمتغير المستقل  $x$  .

فإذا فرضنا أن تجربتنا تضم متغيراً واحداً  $x$  وآخر  $Y$  ، عندئذ تأخذ البيانات شكل الأزواج :  $(x_i, y_i) ; i = 1, 2, \dots, n$  . فإذا خضعت قيم  $x$  للمراقبة ( أى للفحص ) أى أن التجربة قد صممت عندئذ تكون العملية التجريبية هى اختيار القيم  $x_i$  سلفاً ، ومن ثم مراقبة القيم الموافقة  $Y_i$  .

لنفرض أن كل الأوساط  $\mu_{Y_x}$  تقع على مستقيم ( أى أن هذه الأوساط ترتبط ارتباطاً خطياً بـ  $x$  ) ولنفرض أن القيم الملاحظة ( المراقبة ) لـ  $Y$  ستحدد عن هذا الخط المستقيم والذي يمثل هذه العلاقة الخطية ( أى ستقع تحت أو فوق الخط المستقيم ) بمقدار عشوائى  $E$  يسمى بالخطأ ، عندئذ يمكن كتابة هذا المتغير  $Y_i = Y_{x_i} + E_i$  على النحو التالى :

$$Y_i = \mu_{Y|x_i} + E_i = \lambda + \beta x_i + E_i$$

حيث يمثل  $E_i$  متغيراً عشوائياً وسطه حتماً معدوم . وتحقق كل ملاحظة  $(x_i, y_i)$  في العينة العلاقة :

$$y_i = \lambda + \beta x_i + \varepsilon_i .$$

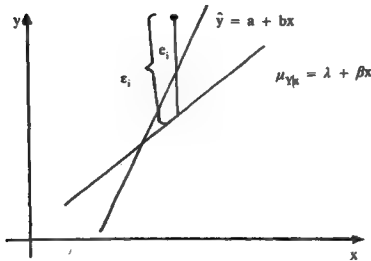
حيث يمثل  $\varepsilon_i$  قيمة يفترضها المتغير  $E_i$  ، وذلك عندما يفترض  $Y_i$  القيمة  $y_i$  ، وبصورة مماثلة وباستخدام خط تقدير الانحدار يمكننا كتابة :

$$\hat{y} = a + bx$$

وتحقق كل ملاحظة من الملاحظات  $(x_i, y_i)$  العلاقة :

$$y_i = a + bx_i + e_i$$

حيث يسمى  $e_i$  بالباقي . يوضح الشكل (٨, ١) الفرق بين  $\varepsilon_i, e_i$



الشكل (٨, ١)

سنحاول إيجاد  $a, b$  ، وهما تقديري  $\beta, \lambda$  بحيث يكون مجموع مربعات البواقي أصغرياً ، نسمى عادة مجموعات مربعات البواقي بمجموع مربعات الأخطاء حول خط الانحدار ونرمز له بالرمز SSE . كما نسمى هذه الطريقة بطريقة المربعات الصغرى ، وبما أننا سنحاول إيجاد كل من  $a, b$  بحيث يكون المقدار التالى أصغرياً :

$$SSE = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - a - bx_i)^2$$

لذلك فإننا سنقوم أولاً باشتقاق SSE جزئياً بالنسبة لـ  $a$  ، ثم بالنسبة لـ  $b$  حيث نجد أن :

$$\frac{\partial(SSE)}{\partial a} = -2 \sum_{i=1}^n (y_i - a - bx_i)$$

$$\frac{\partial(SSE)}{\partial b} = -2 \sum_{i=1}^n (y_i - a - bx_i)x_i$$

وبمساواة كل مشتقة من هاتين المشتقتين الجزئيتين بالصفر ، فإننا نحصل على المعادلتين :

$$na + b \sum_{i=1}^n x_i = \sum_{i=1}^n y_i$$

$$a \sum_{i=1}^n x_i + b \sum_{i=1}^n x_i^2 = \sum_{i=1}^n x_i y_i$$

اللتين تسميان بالمعادلتين الطبيعيين واللتين تقودان إلى :

$$b = \frac{n \sum_{i=1}^n x_i y_i - (\sum_{i=1}^n x_i)(\sum_{i=1}^n y_i)}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2}$$

ومن المعادلة الطبيعية الأولى نجد أن :

$$a = \bar{y} - b\bar{x}$$



Estimating the medians  $\beta, \lambda$ تعريف (٨, ٢) تقديري الوسيطين  $\lambda, \beta$ 

إن تقدير خط الانحدار  $\mu_{y|x} = \lambda + \beta x$  والمحسوب من خلال العينة  $(x_i, y_i) : i = 1, 2, \dots, n$  هو المستقيم :

$$\hat{y} = a + bx$$

حيث يمثل :

$$b = \frac{n \sum_{i=1}^n x_i y_i - \left( \sum_{i=1}^n x_i \right) \left( \sum_{i=1}^n y_i \right)}{n \sum_{i=1}^n x_i^2 - \left( \sum_{i=1}^n x_i \right)^2}$$

$$a = \bar{y} - b\bar{x}$$

مثال (٨, ٣)

لدى مهندس كيميائي عشر قراءات لدرجة الحرارة ، فإذا علمت أن الفحوص التجريبية لدرجة الحرارة قد أعطت :

جدول (٨, ١)

درجة الحرارة $F^\circ$	x	100	110	120	130	140	150	160	170	180	190
درجة الحرارة التي أوضحها التجربة	y	45	51	54	61	66	70	74	78	85	89

ما هي العلاقة الخطية بين هذه البيانات ( أى ما هو انحدار y على x ) ؟

الحل

نلاحظ من شروط المسألة أن :

$$\sum_{i=1}^n x_i = 1450, \sum_{i=1}^{10} y_i = 673, \bar{y} = 67.3$$

$$\sum_{i=1}^n x_i^2 = 218500, \sum_{i=1}^n x_i y_i = 101570, \bar{x} = 145$$

لذلك فإن :

$$b = \frac{10(101570) - (1450)(673)}{10(218500) - (1450)^2} = 0.483$$

$$a = 67.3 - (0.483)(145) = -2.735$$

ومكذا فإن الخط المطلوب هو :

$$\hat{y} = -2.732 + 0.483 x$$

وهو يمثل خط انحدار  $y$  على  $x$ .

مثال (٨, ٤)

أوجد تقديراً لخط انحدار  $x$  على  $y$  للمعلومات التالية :

x	1.5	1.8	2.4	3.0	3.5	3.9	4.4	4.8	5.0
y	4.8	5.7	7.0	8.3	10.9	12.4	13.1	13.6	15.3

الحل

نلاحظ أن :

$$\sum_{i=1}^9 x_i = 30.3, \sum_{i=1}^9 x_i^2 = 115.11, \bar{x} = 3.3667$$

$$\sum_{i=1}^9 y_i = 91.1, \sum_{i=1}^9 x_i y_i = 345.09, \bar{y} = 10.1222$$

وبالتعويض في التعريف (٨,٢) نجد أن :

$$b = \frac{(9)(345.09) - (30.3)(91.1)}{(9)(115.11) - (30.3)^2} = 2.9303$$

$$a = 10.1222 - (2.9303)(3.3667) = 0.2568$$

وهكذا فإن تقدير خط انحدار  $y$  على  $x$  هو :

$$\hat{y} = 0.2568 + 2.9303 x$$

#### (٨,٤) خواص تقديرات المربعات الصغرى

##### Properties of the least squares estimators

لقد فرضنا سابقاً أن الخطأ في العلاقة :

$$y_i = \lambda + \beta x_i + E_i$$

يمثل متغيراً توقعه الرياضي صفراً ، لنفرض إضافة لذلك أن لكل متغير من المتغيرات  $E_i$  توزيعاً طبيعياً بالتباين  $\sigma^2$  ، وأن هذه المتغيرات  $E_1, E_2, \dots, E_n$  مستقلة فيما بينها ، لنفتش عن وسط وتباين تقديري  $\beta, \lambda$  .

يجب ألا يغيب عن أذهاننا بأن قيم  $a, b$  الممثلان لتقديري  $\lambda, \beta$  على الترتيب تحسب من خلال عينة  $n$  من الملاحظات ، وأنه يمكن النظر إلى التقديرات المختلفة لـ  $\lambda, \beta$  ( المحسوبة من عينات مختلفة لها نفس الحجم  $n$  ) كقيم يفترضها المتغيران العشوائيان  $A$  و  $B$  . وحيث أن قيم  $x$  تظل ثابتة ، لذا فإن قيم  $A$  و  $B$  تتعلق باختلافات قيم  $y$  أو بشكل أدق بقيم المتغيرات العشوائية  $Y_1, Y_2, \dots, Y_n$  . إن الفرضيات التوزيعية لـ  $E_i$  تؤكد على

أن  $Y_i$  (حيث  $i = 1, 2, \dots, n$ ) تتوزع بطريقة ماثلة ، وأن لكل منها توزيعاً احتمالياً طبيعياً بالوسط  $\lambda + \beta x_i$  والتباين  $\sigma^2$  وحيث إن التقدير :

$$\begin{aligned} B &= \frac{n \sum_{i=1}^n x_i y_i - \left( \sum_{i=1}^n x_i \right) \left( \sum_{i=1}^n y_i \right)}{n \sum_{i=1}^n x_i^2 - \left( \sum_{i=1}^n x_i \right)^2} \\ &= \frac{\sum_{i=1}^n (x_i - \bar{x}) (y_i - \bar{Y})}{\sum_{i=1}^n (x_i - \bar{x})^2} \\ &= \frac{\sum_{i=1}^n (x_i - \bar{x}) Y_i}{\sum_{i=1}^n (x_i - \bar{x})^2} \end{aligned}$$

يمثل دالة خطية بالمتغيرات العشوائية  $Y_1, Y_2, \dots, Y_n$  والأمثال :

$$a_i = \frac{x_i - \bar{x}}{\sum_{i=1}^n (x_i - \bar{x})^2}, \quad i = 1, 2, \dots, n$$

فإننا نستنتج من النظرية (٥، ١) أن للمتغير  $B$  توزيعاً طبيعياً بالوسط :

$$\begin{aligned} \mu_B = E(B) &= \frac{\sum_{i=1}^n (x_i - \bar{x}) E(Y_i)}{\sum_{i=1}^n (x_i - \bar{x})^2} \\ &= \frac{\sum_{i=1}^n (x_i - \bar{x}) (\lambda + \beta x_i)}{\sum_{i=1}^n (x_i - \bar{x})^2} = \beta \end{aligned}$$

والتباين :

$$\sigma_B^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2 \sigma_{Y_i}^2}{\left[ \sum_{i=1}^n (x_i - \bar{x})^2 \right]^2} = \frac{\sigma^2}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

يمكننا أن نبين أن المتغير العشوائى A هو أيضا متغير عشوائى وسطه :

$$\mu_A = \lambda$$

وتبينه :

$$\sigma_A^2 = \frac{\sum_{i=1}^n x_i^2}{n \sum_{i=1}^n (x_i - \bar{x})^2}$$

ولكى تتمكن من بناء استدلال حول  $\lambda$  و  $\beta$  فإن من الضرورى الوصول إلى تقدير للتباين  $\sigma^2$  الذى يظهر فى عبارتى تباين A و B السابقتين ، ومن المناسب لإيجاد مثل هذا التقدير تعريف الرموز التالية :

$$S_{xx} = \sum_{i=1}^n (x_i - \bar{x})^2 = \sum_{i=1}^n x_i^2 - \frac{(\sum_{i=1}^n x_i)^2}{n}$$

$$S_{yy} = \sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n y_i^2 - \frac{(\sum_{i=1}^n y_i)^2}{n}$$

$$S_{xy} = \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) = \sum_{i=1}^n x_i y_i - \frac{(\sum_{i=1}^n x_i)(\sum_{i=1}^n y_i)}{n}$$

وهكذا نستطيع كتابة مجموع مربعات الأخطاء على النحو التالى :

$$\begin{aligned} SSE &= \sum_{i=1}^n (y_i - a - bx_i)^2 \\ &= \sum_{i=1}^n [(y_i - \bar{y}) - b(x_i - \bar{x})]^2 \\ &= \sum_{i=1}^n (y_i - \bar{y})^2 - 2b \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) + b^2 \sum_{i=1}^n (x_i - \bar{x})^2 \end{aligned}$$

$$= S_{yy} - 2b S_{xy} + b^2 S_{xx}$$

$$= S_{yy} - b S_{xy} ,$$

$$b = \frac{S_{xy}}{S_{xx}} \quad \text{حيث :}$$

يمكننا بعد هذا أن نسوق ونبرهن النظرية التالية :

### نظرية

إن التقدير غير المتحيز لـ  $\sigma^2$  يعطى بالعلاقة :

$$S^2 = \frac{SSE}{n-2} = \frac{S_{yy} - b S_{xy}}{n-2}$$

### البرهان

بتفسير مجموع مربعات الأخطاء كمتغير عشوائى تتغير قيمة بإعادة التجربة مرات عديدة ، عندئذ بإمكاننا كتابة :

$$SSE = S_{yy} - B S_{xx}$$

$$= S_{yy} - B^2 S_{xx}$$

$$(S_{xy} = B S_{xx}) \quad \text{لأن :}$$

$$= \sum_{i=1}^n (y_i - \bar{y})^2 - B^2 \sum_{i=1}^n (x_i - \bar{x})^2$$

$$= \sum_{i=1}^n y_i^2 - n \bar{Y}^2 - \left( \sum_{i=1}^n x_i^2 - n \bar{x}^2 \right) E(B^2)$$

وبأخذ التوقع الرياضى للطرفين نجد أن :

$$\begin{aligned} E(SSE) &= \sum_{i=1}^n E y_i^2 - n E(\bar{Y}^2) \\ &\quad - \left( \sum_{i=1}^n x_i^2 - n \bar{x}^2 \right) E(B^2) \end{aligned}$$

وحيث إن :

$$E y_i^2 = \sigma_{Y_i}^2 + \mu_{Y_i}^2$$

$$E \bar{Y}^2 = \sigma_{\bar{Y}}^2 + \mu_{\bar{Y}}^2$$

$$E(B^2) = \sigma_B^2 + \mu_B^2$$

وبالتبديل فى المعادلة السابقة فإننا نجد أن :

$$\begin{aligned} E(SSE) &= \sum_{i=1}^n (\sigma_{Y_i}^2 + \mu_{Y_i}^2) - n \sigma_{\bar{Y}}^2 + \mu_{\bar{Y}}^2 \\ &\quad - \left( \sum_{i=1}^n x_i^2 - n \bar{x}^2 \right) (\sigma_B^2 + \mu_B^2) \\ &= n \sigma^2 + \sum_{i=1}^n (\lambda + \beta x_i)^2 - n \left[ \frac{\sigma^2}{n} + (\lambda + \beta \bar{x})^2 \right] \\ &\quad - \left( \sum_{i=1}^n x_i^2 - n \bar{x}^2 \right) \left( \frac{\sigma^2}{S_{xx}} + \beta^2 \right) \\ &= (n - 2) \sigma^2 \end{aligned}$$

لذلك فإن :

$$E(S^2) = \frac{E(SSE)}{n-2} = \sigma^2$$

أى إن  $S^2$  يمثل تقديراً غير متحيز لـ  $\sigma^2$ .

### (٨,٥) نهايات الثقة واختبارات المعنوية

#### Confidence limits and tests of significance

لا بد للدارس قبل كل شيء من أن يستقرىء ما إذا كانت هناك أصلاً علاقة بين  $Y$  و  $x$ . وبتعبير آخر هل تقدم المعلومات الإحصائية المتوفرة لديه دليلاً كافياً على وجود علاقة خطية بين  $Y$  و  $x$  فوق فترة معينة لقيم  $x$ ؟ والسؤال المطروح يتعلق بقيمة  $\beta$ .

إن قولنا بأن  $Y$  و  $x$  لا يرتبطان ببعضهما خطياً يكافئ القول بأن  $\beta = 0$ ، وهكذا نخلص إلى القول فى أننا نرغب فى اختبار الفرضية  $\beta = 0$  ضد الفرضية البديلة  $\beta \neq 0$ ، وحيث إن  $\beta$  يمثل متغيراً عشوائياً طبيعياً وأن  $\frac{(n-2)S^2}{\sigma^2}$  يمثل متغيراً فى نوع كائى - مربع بـ  $(n-2)$  درجة من الحرية لذلك وحسب النظرية (٥,٧) فإن المتغير :

$$T = \frac{(B - \beta) / (\sigma / \sqrt{S_{xx}})}{S / \sigma} = \frac{B - \beta}{S / \sqrt{S_{xx}}}$$

يتوزع وفقاً للتوزيع  $t$  بـ  $(n-2)$  درجة من الحرية يمكننا استخدام الإحصاء  $T$  لاجتياز 100%  $(1 - \alpha)$  فترة ثقة للمعامل  $\beta$  فى معادلة خط الانحدار  $\mu_{Yx} = \lambda + \beta x$  هى الفترة :

$$b - \frac{t_{\alpha/2} \cdot S}{\sqrt{S_{xx}}} < \beta < b + \frac{t_{\alpha/2} \cdot S}{\sqrt{S_{xx}}}$$

حيث تمثل  $t_{\alpha/2}$  قيمة الإحصاء  $t$  بـ  $(n-2)$  درجة من الحرية .



## مثال (٨،٥)

أجريت تجربة صندوق القص لتحديد العلاقة بين الإجهاد المتعامد (normal stress) وصامد القص (Shear resistance) للتربة فأعطت النتائج التالية :

x	الإجهاد المتعامد (KN/m <sup>2</sup> )	11	13	15	17	19	21
y	صامد القص (KN/m <sup>2</sup> )	15.2	17.7	19.3	21.5	23.9	25.4

أوجد في معادلة الانحدار  $\mu_{y|x} = \lambda + \beta x$  فترة ثقة للمعامل  $\beta$  .

الحل

نلاحظ أن :

$$\sum_{i=1}^6 x_i = 96, \sum_{i=1}^6 y_i = 123, \sum_{i=1}^6 y_i^2 = 2595.44$$

$$\sum_{i=1}^6 x_i^2 = 1606, \sum_{i=1}^6 x_i y_i = 2039.8$$

لذلك فإن :

$$S_{xx} = 1606 - \frac{(96)^2}{6} = 70$$

$$S_{yy} = 2595.44 - \frac{(123)^2}{6} = 73.94$$

$$S_{xy} = 2039.8 - \frac{(96)(123)}{6} = 71.8$$

وهذا يعنى بأن  $b = 1.0257$  وهكذا فإن :

$$S^2 = \frac{S_{yy} - b S_{xy}}{n - 2}$$

$$= \frac{73.94 - (1.0257)(71.8)}{4} = 0.073685$$

وبأخذ الجذر التربيعي نجد أن  $\sqrt{S_{xx}} = 8.3666$  . كما نجد أيضا أن  $s = 0.27144$  ، من الجدول  $v$  نجد أن  $t_{0.025} = 2.776$  وذلك من أجل 4 درجات حرية . وهكذا فإن 95% فترة الثقة المطلوبة لـ  $\beta$  هي :

$$\left( 1.0257 - \frac{(2.776)(0.073685)}{8.3666}, 1.0257 + \frac{(2.776)(0.073685)}{8.3666} \right)$$

أو : (1.00126 , 1.050148)

أي أن :  $1.00125 < \beta < 1.050148$

لاختيار الفرضية  $H_0: \beta = B_0$  ضد فرضية بديلة مناسبة ، فإننا نستخدم أيضا التوزيع  $t$  بـ  $(n - 2)$  درجة من الحرية ، وذلك لتجديد المنطقة الحرجة وبعد ذلك نبني قرارنا على القيمة :

$$t = \frac{b - B_0}{s / \sqrt{S_{xx}}}$$

مثال (٨,٦)

باستخدام التقدير  $b = 1.0257$  في المثال السابق اختبر الفرضية  $\beta = 0.98$  ضد الفرضية  $\beta > 0.98$  ، استخدم مستوى للمعنوية 0.01 5

الحل

نلاحظ أن :

$$H_0: \beta = 0.98$$

(١) فرض العدم هو :

(٢) الفرضية البديلة هي :  $H_1: \beta > 0.98$

(٣) أن مستوى المعنوية هو :  $\alpha = 0.01$

(٤) المنطقة الحرجة هي :  $T > 3.747$

(٥) الإحصاء المحسوب من العينة هو :  $T = \frac{1.0257 - 0.98}{0.073685 / 8.3666} = 5.189$

(٦) الاستنتاج : نرفض فرض العدم  $H_0: \beta = 0.98$  ونقبل الفرضية البديلة  $H_1: \beta > 0.98$  وذلك لأن قيمة الإحصاء  $t = 5.189$  وقعت داخل المنطقة الحرجة  $T > 3.747$ .

يمكن بنفس الطريقة أن نبحث عن فترة الثقة ، وأن ندرس اختبار الثقة للمعامل  $\lambda$  إذا أخذنا بعين الاعتبار أن المتغير  $A$  طبيعي . من السهل التأكد من أن للإحصاء :

$$T = \frac{A - \lambda}{S \sqrt{\frac{n}{\sum_{i=1}^n x_i^2} / n S_{xx}}}$$

توزيعاً من نوع  $t$  بـ  $(n - 2)$  درجة من الحرية ، وبهذا الشكل نستنتج أن  $100\% (1 - \alpha)$  فترة الثقة للمعامل  $\lambda$  في خط الانحدار  $\mu_{y|x} = \lambda + \beta x$  هي :

$$a = \frac{t_{\alpha/2} \cdot S \sqrt{\frac{n}{\sum_{i=1}^n x_i^2}}}{\sqrt{n \cdot S_{xx}}} < \lambda < a + \frac{t_{\alpha/2} \cdot S \sqrt{\frac{n}{\sum_{i=1}^n x_i^2}}}{\sqrt{n \cdot S_{xx}}}$$

حيث يمثل  $t_{\alpha/2}$  قيمة  $t$  بـ  $(n - 2)$  درجة من الحرية .

لاختبار فرضية  $\lambda_0: \lambda = \lambda_0$  ضد فرضية بديلة مناسبة ، يمكننا استخدام التوزيع  $t$  بـ  $(n - 2)$  درجة حرية ، وذلك لتحديد المنطقة الحرجة ، وبعد ذلك نبني قرارنا حول القيمة :

$$t = \frac{\lambda - \lambda_0}{S \sqrt{\frac{n}{\sum_{i=1}^n x_i^2} / n S_{xx}}}$$

لنفرض أننا نهم بتوقع  $Y$  من أجل القيمة  $x = x_0$  عندئذ يمثل المقدار :

$$Y_0 = A + B x_0$$

$$\mu_{Y_{x_0}} = \lambda + \beta x_0 \quad \text{تقديرًا للوسط}$$

$$E(\hat{Y}_0) = E(A + B x_0) = \lambda + \beta x_0 \quad \text{وحيث إن :}$$

فهذا يعنى أن التقدير  $\hat{Y}_0$  هو تقدير غير متحيز للوسط  $\mu_{Y_{x_0}}$

بالتباين :

$$\sigma_{\hat{Y}_0}^2 = \sigma_A^2 + \beta x_0 = \sigma_{\hat{Y}}^2 + B(x_0 - \bar{x})$$

$$= \sigma^2 \left[ \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}} \right]$$

وهكذا فإن بإمكاننا إنشاء 100% (1 -  $\alpha$ ) فترة ثقة للوسط  $\mu_{Y_{x_0}}$  بواسطة الإحصاء :

$$T = \frac{\hat{Y}_0 - \mu_{Y_{x_0}}}{S \sqrt{\frac{1}{n} + [(x_0 - \bar{x})^2 / S_{xx}]}}$$

والذى يتوزع وفق توزيع  $t$  بـ (n - 2) درجة من الحرية .

إن 100% (1 -  $\alpha$ ) فترة الثقة للوسط  $\mu_{Y_{x_0}}$  هى :

$$\hat{y}_0 - t_{\alpha/2} S \sqrt{\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}}} < \mu_{Y_{x_0}} < \hat{y}_0 + t_{\alpha/2} S \sqrt{\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}}}$$

حيث تمثل  $t_{\alpha/2}$  قيمة التوزيع  $t$  بـ (n - 2) درجة من الحرية .

مثال (٨,٧)

أوجد فى المثال (٨,٥) 95% فترة للوسط  $\mu_{Y_{x_0}}$  أى وسط الصامد وذلك من أجل

$$x = 25 \text{ kN/m}^2$$

## الحل

نلاحظ من معادلة الانحدار أنه من أجل  $x = 25$  فإن تقدير الوسط  $\hat{y}_0$  هو :

$$\hat{y}_0 = 4.0888 + (1.0257)(25) = 29.73 \text{ kN/m}^2$$

إن 95% فترة ثقة لهذا التقدير هي :

$$\left( \hat{y}_0 - t_{0.025} S \sqrt{\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}}}, \hat{y}_0 + t_{0.025} S \sqrt{\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}}} \right)$$

أى :

$$\left( (29.73 - (2.776)(0.27144)) \sqrt{\frac{1}{6} + \frac{(25 - 16)^2}{70}}, \right.$$

$$\left. 29.739 + (2.776)(0.27144) \sqrt{\frac{1}{6} + \frac{(25 - 16)^2}{70}} \right)$$

$$= (28.863, 30.606)$$

$$28.863 < \mu_{y|x} < 30.606 \quad \text{أى :}$$

نفرض أننا نريد البدء بالصفات التوزيعية للفروقات بين الترتيبات المحسوبة من معادلة خط الانحدار في العينات المتكررة والترتيب الحقيقي المقابل  $y_0$  عند القيمة  $x = x_0$  ، عندئذ يمكننا التفكير بالفرق  $\hat{y}_0 - y_0$  كقيمة للمتغير العشوائى  $\hat{Y}_0 - Y_0$  والذي يمكن البرهان بأن له توزيع طبيعى بالوسط :

$$\mu_{\hat{Y}_0 - Y_0} = E(\hat{Y}_0 - Y_0)$$

$$= E[(A + Bx_0) - (\lambda + \beta x_0 + E_0)]$$

$$= 0$$

والتباين :

$$\begin{aligned}\sigma_{\hat{Y}_0}^2 - Y_0 &= \sigma_A^2 + Bx_0 - E_0 \\ &= \sigma_y^2 + B(x_0 - \bar{x}) - E_0 \\ &= \sigma^2 \left[ 1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}} \right]\end{aligned}$$

وعندئذ يمكن الحصول على 100% (1 - α) فترة ثقة للقيمة الوحيدة المتنبئة  $y_0$  من الإحصاء :

$$T = \frac{\hat{Y}_0 - Y_0}{S \sqrt{1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}}}}$$

والذى يتوزع وفقاً للتوزيع  $t$  بـ (n - 2) درجة من الحرية . وهكذا فإن 100% (1 - α) فترة الثقة للقيمة الوحيدة  $y_0$  هي :

$$\hat{y}_0 - t_{\alpha/2} S \sqrt{1 + \frac{1}{n} + \frac{(X_0 - \bar{X})^2}{S_{xx}}} < y_0 < \hat{y}_0 + t_{\alpha/2} S \sqrt{1 + \frac{1}{n} + \frac{(X_0 - \bar{X})^2}{S_{xx}}}$$

حيث يمثل  $t_{\alpha/2}$  قيمة  $t$  بـ (n - 2) درجة من الحرية .

مثال (٨,٨)

أوجد في المثال (٨,٧) 95% فترة للقيمة التقديرية الوحيدة لصامد المقص والموافقة لـ  $x = 25 \text{ KN/m}^2$

الحل

لدينا  $n = 6$ ,  $x_0 = 25$ ,  $\bar{x} = 16$ ,  $\hat{y}_0 = 29.739$ ,  $S_{xx} = 70$ ,  $S = 0.27144$ , وأخيراً

$t_{0.025} = 2.776$  من أجل 4 درجات حرية ، لذلك فإن الثقة لـ  $y_0$  هي :

$$29.739 \pm (2.776) (0.27144) \sqrt{1 + \frac{1}{6} + \frac{(25 - 16)^2}{70}}$$

$$29.739 \pm 1.1487$$

$$(20.59 , 30.888) \text{ KN/m}^2$$

أى :

### ٨,٦) تحليل التباين Analysis of variance

تُعالج مسألة تحليل نوعية تقدير خط الانحدار غالباً من خلال إجراء يسمى بتحليل التباين ، حيث تجزأ وفق هذا الإجراء الاختلافات الكلية في المتغير المرتبط  $y$  إلى عوامل تُراقب وتُعالج بشكل نظامي .

لنفرض أن لدينا  $n$  من المعلومات التجريبية على شكل أزواج  $(x_i, y_i)$  ، ولنفرض أيضاً أنه قد تم تقدير خط الانحدار ، لقد أوضحنا في الفقرة (٨,٣) عند تقديرنا لـ  $\sigma^2$  أن :

$$S_{yy} = b S_{xy} + SSE$$

وهذا يعنى أننا جزأنا مجموع المربعات الكلى في  $y$  إلى جزأين سنرمز لهذين الجزأين بالرمزين :

$$SST = S_{yy} , \quad SSR = b S_{xy}$$

وعندئذ ستأخذ المتطابقة السابقة الشكل الجديد التالى :

$$SST = SSE + SSR$$

نسعى الجزء الأول في الطرف الأيمن من العلاقة السابقة  $SSR$  بمجموع المربعات العائد لخط الانحدار ، وهذا المجموع يؤثر على مقدار الانحراف في قيم  $y$  والموضحة بالنظام ( في هذه الحالة الخط المستقيم المسلم به ) ، أما الجزء الثانى  $SSE$  فهو يمثل مجموع مربعات

الأخطاء ، والذي يؤثر على الاختلافات حول خط الانحدار ، وحيث إن  $\frac{SSE}{\sigma^2}$  ،  $\frac{SSR}{\sigma^2}$  لتغيرين من نوع كاي - مربع بـ (1) ، (n - 2) درجة حرة على الترتيب لذا فإن تمثل قيمة لتغير عشوائى من نوع كاي - مربع بـ (n - 1) درجة حرية . وعند اختبار فرضية  $H_0: \beta = 0$  ضد فرضية بديلة  $H_1: \beta \neq 0$  حيث تشير فرضية العدم إلى أن الاختلاف في Y لا يُفسر بواسطة الخط المستقيم ، وإنما بواسطة التردد العشوائى-أو بواسطة الخط فإننا نحسب :

$$f = \frac{SSR / 1}{SST / (n - 2)} = \frac{SSR}{s^2}$$

ثم نرفض  $H_0$  عند مستوى المعنوية  $\alpha$  إذا كان  $f > f_{\alpha}(1, n - 2)$  . ونستنتج أن هناك مقدار اختلاف هام في الإجابة المحسوبة بواسطة النظام المسلم به ( أى دالة الخط المستقيم ) . أما إذا وقع الإحصاء F في منطقة القبول ، فإننا نستنتج أن المعلومات لا تؤثر بشكل كاف في دعم النظام المسلم به .

وعند حل هذه المسألة نبدأ بحساب :

$$SST = S_{yy}$$

$$SSR = bS_{xy}$$

وبعد ذلك نستخدم المتطابقة :

$$SSE = SST - SSR$$

ثم نقوم بتلخيص هذا في جدول كالتالى ، ( وهو ما يسمى عادة بجدول تحليل التباين ) :

مصدر الاختلاف	مجموع المربعات	درجة الحرية	وسط التوزيع	f المحسوبة
الانحدار	SSR	1	SSR	$\frac{SSR}{S^2}$
الخطأ	SSE	n - 2	$S^2 = \frac{SSE}{n - 2}$	
المجموع	SST	n - 1		

جدول (٨،١) أو جدول تحليل التباين من أجل  $\beta = 0$



## ملاحظة

لقد استخدمنا في الفقرة (٨،٤) عند اختبار الفرضية  $H_0: \beta = \beta_0$  ضد الفرضية البديلة  $H_1: \beta \neq \beta_0$  الإحصاء :

$$T = \frac{B - \beta_0}{S / \sqrt{S_{xx}}}$$

حيث يتوزع  $T$  وفق توزيع  $t$  بـ  $(n - 2)$  درجة حرية ، ولقد فرضنا فرض العدم  $H_0$  عندما وقع  $|T| > t_{\alpha/2}$  عند مستوى المعنوية  $\alpha$  ، ومن المهم الإشارة إلى أنه في الحالة التي تكون فيها  $H_0: \beta = 0$  ، فإن قيمة الإحصاء  $T$  تصبح :

$$t = \frac{b}{S / \sqrt{S_{xx}}}$$

وهي تتطابق مع تلك المذكورة في الجلول (٨،١) ، أعنى أن فرض العدم يقرر الاختلاف في  $y$  ناشئ عن الحظ . ويستخدم تحليل التباين التوزيع  $F$  أكثر من التوزيع  $t$  . غير أن الإجرائين متماثلان ويبدو هذا بوضوح بكتابة :

$$t^2 = \frac{b^2 S_{xx}}{S^2} = \frac{b S_{xy}}{S^2} = \frac{SSR}{S^2}$$

والمطابقة لقيمة  $f$  المستخدمة في تحليل التباين . إن العلاقة الأساسية بين التوزيع  $t$  بـ  $v$  درجة من الحرية ، والتوزيع  $F$  بـ  $1$  و  $v$  درجات حرية يعطى بالعلاقة :

$$t_{\alpha/2}^2 = f_{\alpha}(1, v)$$

(٨،٧) القياسات المتكررة لـ  $y$ The frequent measurement for  $y$ 

من أجل الحصول على معلومات كمية تتعلق بملائمة النظام المستخدم ، فإن على المجرّب أن يجري ملاحظات متكررة لكل قيمة لـ  $x$  ، بينما لا يهم حدوث مثل هذه التكرارات عند تقدير كل من  $\beta$  و  $\beta_0$  ، وبالحقيقة فإن وجود الملاحظات المتكررة تحت

تصرف المجرّب تمكنه من إجراء اختبار أهمية يهدف إلى تحديد ما إذا كان النظام المستخدم ملائماً أم لا .

لنختار عينة عشوائية مؤلفة من  $n$  ملاحظة ، وذلك باستخدام  $k$  قيمة مختلفة لـ  $x$  مثلاً  $x_1, x_2, \dots, x_k$  ، وبحيث تحوى العينة  $n_1$  ملاحظة للمتغير  $Y_1$  الموافق لـ  $x_1$  ،  $n_2$  ملاحظة للمتغير  $Y_2$  الموافق لـ  $x_2, \dots, n_k$  ملاحظة للمتغير  $Y_k$  الموافق لـ  $x_k$  ، وبحيث يكون  $n = \sum_{i=1}^k n_i$  . لنفرض أن  $y_{ij}$  تمثل القيمة ذات الرقم  $j$  للمتغير العشوائى  $Y_i$  ، ولنفرض أن :

$$y_i = T_i = \sum_{j=1}^{n_i} y_{ij}$$

$$\bar{y}_i = \frac{T_i}{n_i}$$

لنفرض مثلاً أن  $n_4 = 3$  عندئذ تكون قياسات  $Y_4$  الموافقة لـ  $x = x_4$  هى  $y_{41}, y_{42}, y_{43}$  وسيكون أيضاً :

$$T_4 = y_{41} + y_{42} + y_{43}$$

يتكون مجموع مربعات الأخطاء من جزئين يمثل الأول منهما الكمية الناشئة عن الاختلاف بين قيم  $Y$  والقيم المعطاة لـ  $x$  ، ويسمى هذا الجزء بنقصان تطابق الملائمة (lack of fit contribution) ، وهو يعكس اختلافاً عشوائياً مجرداً أو خطأً تجريبياً صرفاً ، بينما يمثل الجزء الثانى قياس للاختلاف النظامى الذى يحدث تحت شروط عالية الدرجة ، وفى حالتنا هذه فإن هذه الشروط فى  $x$  غير التداخل الحطى أو التداخل من الدرجة الأولى . نلاحظ أنه باختيارنا لنظام خطى ما ، فإننا نفرض أساساً أن الجزء الثانى غير موجود ولذلك فإن مجموع مربعات الأخطاء تنسب إلى خطأ عشوائى . وفى هذه الحالة

يكون المقدار  $SSE = \frac{SSE}{n-2}$  ممثلاً لتقدير ما غير متحيز لـ  $\sigma^2$  . ومع ذلك إذا كان النظام غير منسجم بشكل كاف مع المعلومات فعندئذ يكون مجموع مربعات الأخطاء متضخماً وتحدث تقديراً غير متحيز لـ  $\sigma^2$  ، وسواء أكان النظام ملائماً للمعلومات أو غير ملائم ، فإن بإمكان المجرّب دائماً الحصول على تقدير غير متحيز لـ  $\sigma^2$  ، وذلك عندما تتكرر الملاحظات .

كما نلاحظ أن :

$$S_1^2 = \sum_{j=1}^n (y_{ij} - \bar{y})^2, i = 1, 2, \dots, k$$

وأنه من أجل كل القيم المختلفة الـ  $k$  لـ  $x$  فإن :

$$S^2 = \frac{\sum_{i=1}^k (n_i - 1) S_i^2}{n - k} = \frac{\sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y})^2}{n - k}$$

يمثل بسط  $S^2$  قياساً للخطأ التجريبي الصرف ، هذا ويمكن فصل مجموع مربعات الأخطاء بطريقة حسابية إلى جزئين يمثل أحدهما الخطأ التجريبي الصرف ، كما يمثل الثاني النقطان في الملائمة ، (lack of fit) وذلك على النحو التالي :

١ — نبدأ بحساب مجموع مربعات الخطأ الصرف :

$$\sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y})^2 = \sum_{i=1}^k \sum_{j=1}^{n_i} y_{ij}^2 - \sum_{i=1}^k \frac{T_i^2}{n_i}$$

ولمجموع المربعات السابق  $n - k$  درجة حرية مرتبطة به . أما وسطه فهو يمثل تقديراً غير المتحيز  $S^2 > \sigma^2$  .

٢ — يطرح مجموع مربعات الخطأ الصرف من مجموع مربعات الأخطاء SSE ، فإننا نحصل على مجموع المربعات المعزوة إلى النقصان في الملائمة . هذا ويمكن الحصول على درجات حرية نقصان الملائمة بطرح بسيط  $k - 2 = (n - k) - (n - 2)$  .

يمكن إجمال الحسابات المطلوبة لاختبار فرضية بواسطة القياسات المتكررة لـ  $y$  في مسألة الانحدار بجدول كالتالي :

إن مفهوم النقصان في الملائمة مهم جداً في تطبيقات تحليل الانحدار . وفي الحقيقة فإن الحاجة إلى تصميم تجربة تقدم بياناً عن النقصان في الملائمة تصبح أكثر من مسألة ، وتصبح التقنية الأساسية اللازمة لذلك أكثر تعقيداً وبالتأكيد لا يمكن للمعرب أن يكون متأكداً دائماً من أن البناء المسلم به ( في هذه الحالة نظام الانحدار الخطي ) صحيح أو حتى يمثل

كافيا . يوضح المثال التالى كيف أن مجموع مربعات الأخطاء ينقسم إلى جزأين يمثلان الخطأ الصرف والنقصان فى الملاءمة .

جدول (٨,٢) تحليل التباين من أجل القياسات المتكررة لـ  $y$

مصدر الاختلاف	مجموع المربعات	درجات الحرية	وسط المربعات	الحساب
الانحدار	SSR	1	SSR	$\frac{SSR}{2}$
الخطأ	SSE	$n - 2$		
النقصان فى الملاءمة	(الصرف - SSE) SEE	$k - 2$	(الصرف - SSE) SSE	$SSE - (SSE \text{ الصرف})$
الخطأ الصرف	الصرف SSE	$n - k$	$\frac{k - 2}{n - k} \text{ الصرف SSE}$	$S^2(k - 2)$
المجموع	SST	$n - 1$		

مثال (٨,٩)

بفرض أن كميات المعادن المستخرجة من مادة خاصة لدى تعرضها لفترات تنشيف ذات أطوال مختلفة هى كالتالى :

الجدول (٨,٣)

$x$ بالساعات	$y$ بالقرامات	
4.4	13.1	14.2
4.5	09.0	11.5
4.8	10.4	11.5
5.5	13.8	14.8
5.7	12.7	15.1
5.9	09.9	12.7
6.3	13.8	16.5
6.9	16.4	15.7
7.5	17.6	16.9
7.8	18.3	17.2

- ١ — فتش عن تقدير للنظام الخطى  $\mu_{yk} = \lambda + \beta x$   
 ٢ — اختيار النقصان فى الملائمة .

## الحل

لدينا  $n_1 = n_2 = \dots = n_{10} = 2$  لذلك فإن :

$$\begin{aligned}
 S_{yy} &= \sum_{i=1}^{10} \sum_{j=1}^2 y_{ij}^2 - \frac{(\sum_{i=1}^{10} \sum_{j=1}^2 y_{ij})^2}{20} \\
 &= 4089.23 - \frac{(281.1)^2}{20} \\
 &= 4089.23 - 3950.8605 \\
 &= 138.3695 \\
 S_{xx} &= \sum_{i=1}^{10} n_i x_i^2 - \frac{(\sum_{i=1}^{10} n_i x_i)^2}{20} \\
 &= 2 (364.59) - \frac{(118.6)^2}{20} \\
 &= 729.18 - 703.298 \\
 &= 25.882 \\
 S_{xy} &= \sum_{i=1}^{10} \sum_{j=1}^2 x_i y_{ij} - \frac{(\sum_{i=1}^{10} n_i x_i) (\sum_{i=1}^{10} \sum_{j=1}^2 y_{ij})}{20} \\
 &= 1714.62 - \frac{(118.6)(281.1)}{20} \\
 &= 47.697
 \end{aligned}$$

$$\bar{y} = 14.055, \bar{x} = 5.93$$

إن معاملي الانحدار هما :

$$b = \frac{S_{xy}}{S_{xx}} = \frac{47.697}{25.882} = 1.8429$$

ومنه :

$$\begin{aligned} a &= \bar{y} - b\bar{x} \\ &= 14.055 - (1.8429)(5.93) \\ &= 3.1266 \end{aligned}$$

لذلك فإن تقدير خط الانحدار هو :

$$\hat{y} = a + bx = 3.1266 + 1.8429x$$

نتبع الطرق الاعتيادية لاختبار نقصان الملائمة حيث نلاحظ أن :

- ١ — فرض العدم : الانحدار في  $x$  خطي :  $H_0$
- ٢ — الفرض البديل : الانحدار في  $x$  غير خطي :  $H_1$
- ٣ — لنفرض أن مستوى المعنوية هو  $\alpha = 0.05$
- ٤ — المنطقة الحرجة هي  $F > 3.07$  بـ 10 درجات حرية .

٥ — الحسابات : لدينا

$$SST = S_{yy} = 138.3695$$

$$SSR = b S_{xy} = (1.8429)(47.697) = 87.9$$

$$SSE = S_{yy} - b S_{xy} = 138.3695 - 87.9 = 50.47$$

ولحساب مجموع مربعات الخطأ الصرف فإننا نكتب أولاً :

$x_1 = 4.4$	$T_1 = 27.3$
$x_2 = 4.5$	$T_2 = 20.5$
$x_3 = 4.8$	$T_3 = 21.9$
$x_4 = 5.5$	$T_4 = 28.6$
$x_5 = 5.7$	$T_5 = 27.8$
$x_6 = 5.9$	$T_6 = 27.6$
$x_7 = 6.3$	$T_7 = 30.3$
$x_8 = 6.9$	$T_8 = 32.1$
$x_9 = 7.5$	$T_9 = 34.5$
$x_{10} = 7.8$	$T_{10} = 35.5$

لذلك فإن :

$$\begin{aligned}
 (\text{SSE الصرف}) &= 4089.23 - \frac{(27.3)^2 + (20.5)^2 + (21.9)^2 + \dots + (35.5)^2}{2} \\
 &= 4089.23 - 4072.86 = 16.375
 \end{aligned}$$

يمكن إجمال النتائج السابقة بجدول كالآتي :

جدول (٨,٦) تحليل التباين في فترات التنشيف المختلفة

مصدر الاختلاف	مجموع المربعات	درجات الحرية	وسط المجموع	حساب f
الانحدار الخطأ نقصان الملائمة الخطأ الصرف	87.900	01	87.9000	53.68
	50.470	18		
	34.095	08	04.2619	02.60
	16.375	10	01.6375	
المجموع	138.37	19	93.7994	56.28

٦ — الاستنتاج : إن تجزئة مجموع الاختلافات الكلية بهذه الطريقة يُظهر اختلافاً مهماً مرتبطاً بالنظام الخطي ، كما تُظهر كمية من الاختلافات غير المهمة والمنسوبة إلى النقصان في الملائمة ، وهكذا فإن المعلومات التجريبية لا تقترح الحاجة إلى افتراض شروط أعلى من الدرجة الأولى في النظام ، ونقبل بذلك فرض العدم .

### (٨,٨) الارتباط The correlation

لقد فرضنا حتى الآن أننا نقوم بملاحظة المتغير المستقل  $x$  ، ولذلك فهذا المتغير ليس عشوائياً ، وفي الحقيقة فإنه يسمى في هذا السياق بمتغير رياضي . وهو قياس في العملية المعنية بخطأ مهمل . ومن الأجلر في كثير من تطبيقات الانحدار التطبيقية الفرض بأن كلا من  $X, Y$  يمثل متغيراً عشوائياً ، وأن القياسات  $(x_i, y_i) : i = 1, 2, \dots, n$  تمثل ملاحظات من دالة الكثافة المشتركة  $f(x, y)$  . ففي علم الآثار مثلاً يمكن الفرض بأن قياسين لنوع معين من عظام جسم كائن بالغ هما متغيران عشوائيان يتبعان توزيعاً ذي بعدين .

يفترض عادة أن التوزيع الشرطي  $f(y|x)$  للمتغير  $Y$  من أجل قيم ثابتة لـ  $X$  هو توزيع طبيعي بالوسط  $\mu_{y|x} = \lambda + \beta x$  ، والتباين  $\sigma_{y|x}^2 = \sigma^2$  ، وأن  $X$  يتوزع بطريقة مماثلة وفقاً للتوزيع الطبيعي بالوسط  $\mu_x$  ، والتباين  $\sigma_x^2$  ، وأن الكثافة المشتركة لـ  $X$  و  $Y$  تعطى بالعلاقة :

$$f(x, y) = n(y|x : \lambda + \beta x, \sigma) \cdot n(x : \mu_x, \sigma_x)$$

$$= \frac{1}{2\pi\sigma_x\sigma} \exp \left\{ -\left(\frac{1}{2}\right) \left[ \left(\frac{y - (\lambda + \beta x)}{\sigma}\right)^2 + \left(\frac{x - \mu_x}{\sigma_x}\right)^2 \right] \right\}$$

حيث إن  $-\infty < x < \infty$  و  $-\infty < y < \infty$  ويكتابة المتغير  $Y$  على النحو :

$$Y = \lambda + \beta X + E$$

حيث يمثل  $X$  هنا متغيراً عشوائياً مستقلاً عن الخطأ العشوائي  $E$  ، فإننا نجد أن :

$$\mu_y = \lambda + \beta \mu_x$$

$$\sigma_y^2 = \sigma^2 + \beta^2 \sigma_x^2$$



وبالتبديل في عبارة  $f(x, y)$  السابقة ، فإننا نحصل على توزيع طبيعي ذاتى بعدين :

$$f(x, y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \exp \frac{-1}{2(1-\rho^2)} \left[ \left( \frac{X-\mu_x}{\sigma_x} \right)^2 - 2\rho \left( \frac{X-\mu_x}{\sigma_x} \right) \left( \frac{Y-\mu_y}{\sigma_y} \right) + \left( \frac{Y-\mu_y}{\sigma_y} \right)^2 \right]$$

$$\rho = 1 - \frac{\sigma_y^2}{\sigma_x^2} = \beta \frac{\sigma_x^2}{\sigma_y^2} , -\infty < y < \infty , -\infty < x < \infty \quad \text{حيث إن}$$

يسمى الثابت  $\rho$  بمعامل الارتباط . وهو يلعب دوراً هاماً في عدد من مسائل تحليل البيانات ذات البعدين . ومن المهم للقارئ أن يفهم التفسير الفيزيائي لمعامل الارتباط والاختلاف بين الارتباط والانحدار . ويظل لتعبير الانحدار هنا معنى ، وبالحقيقة فإن الخط المستقيم  $\mu_{Y|x} = \lambda + \beta X$  يظل اسمه خط انحدار كما في السابق وأن تقديرات  $\lambda, \beta$  تطابق التقديرات الواردة في الفقرة (٨,٢) ، ويأخذ  $\rho$  القيمة صفر عندما يكون  $\beta = 0$  ، ويحدث هذا عندما لا يكون هناك انحداراً خطياً ، وهذا يعنى أن خطأ الانحدار أفقى ، وأن معرفتنا بـ  $X$  لا تفيد في التنبؤ بقيمة  $Y$  . وبما أن  $\sigma_y^2 = \sigma^2$  لذا فإننا نجد أن  $-1 \leq \rho \leq 1$  ، ويأخذ  $\rho$  القيمتين  $\pm 1$  عندما يكون  $\sigma^2 = 0$  . وهذا يعنى وجود علاقة خطية كاملة بين المتغيرين ، وهكذا فإن القيمة  $\rho = 1$  تدل ضمناً على وجود علاقة خطية كاملة بميل موجب ، بينما تشير القيمة  $\rho = -1$  على وجود علاقة خطية كاملة بميل سالب .

هذا ويمكن القول بأن التقديرات العينية لـ  $\rho$  القريبة من الواحد تقتضى إما ارتباط جيداً أو علاقة خطية بين  $X$  و  $Y$  ، بينما تشير القيم القريبة من الصفر إما إلى ارتباط بسيط أو عدمه .

يجب التأكيد على أن قيمة  $\rho$  المحسوبة تقيس في ( أية حالة ) درجة العلاقة بالنسبة لنوع المعادلة المفروضة . فإذا فرضنا معادلة خطية ونتج أن قيمة  $\rho$  تقترب من الصفر ، فهذا يعنى أنه لا يوجد تقريباً علاقة خطية بين المتغيرين ، ولكن هذا لا يعنى أنه لا يوجد

علاقة بين المتغيرين على الإطلاق ، حيث إنه قد يكون هناك بالفعل علاقة كبيرة غير خطية بين هذين المتغيرين . وبصورة أخرى فإن معامل الارتباط يقيس مدى جودة توفيق المعادلة المفترضة للبيانات . إن مصطلح معامل الارتباط يستخدم ليعنى الارتباط الخطى مالم نُشير إلى خلاف ذلك . ويجب إيضاح أن وجود معامل ارتباط مرتفع ، أى يقترب من 1 أو -1 لا يعنى وجود علاقة تبعية مباشرة بين المتغيرين ، فقد يكون هناك معامل ارتباط مرتفع بين عدد الذين تزوجوا وعدد الأشجار المثمرة ، مثل هذه الأمثلة يشار إليها بأنها ارتباط لا معنى له أو ارتباط زائف .

للحصول على تقدير عيني لمعامل الارتباط  $r$  فإننا نعود قليلا إلى الفقرة (٨،٣) لنجد أن مجموع مربعات الأخطاء تعطى بالعلاقة :

$$SSE = S_{yy} - b S_{xy}$$

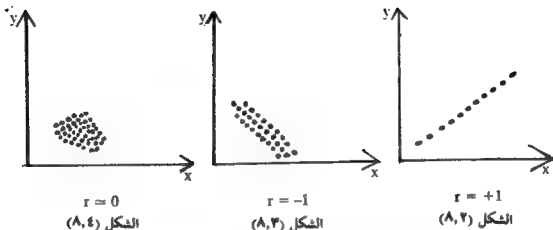
وبقسمة طرفي المعادلة الأخيرة على  $S_{yy}$  وبتبديل  $b = \frac{S_{xy}}{S_{xx}}$  فإننا نحصل على العلاقة :

$$b^2 \frac{S_{xx}}{S_{yy}} = 1 - \frac{SSE}{S_{yy}}$$

إن قيمة  $b^2 \frac{S_{xx}}{S_{yy}}$  تساوى الصفر عندما يكون  $b = 0$  و يحدث هذا عندما لا يكون هناك علاقة خطية بين نقاط العينة . وبما أن  $S_{yy} \geq SSE$  فإننا نستنتج أن  $b^2 \frac{S_{xx}}{S_{yy}}$  تقع بين الصفر والواحد ، لذلك فإن  $b/\sqrt{S_{xx}/S_{yy}}$  تتردد بين -1 ، +1 ، وتوافق القيم السالبة مستقيمات بميول سالبة ، كما توافق القيم الموجبة مستقيمات بميول موجبة ، وسيأخذ هذا المقدار القيمة 1 أو -1 عندما يكون  $SSE = 0$  . وتوضع في هذه الحالة جميع قيم العينة على خط مستقيم . لذلك فإن العلاقة الخطية الكاملة تحدث بين  $X$  و  $Y$  عندما يكون :

$$b/\sqrt{S_{xx}/S_{yy}} = \pm 1$$

تمثل الأشكال (٨،٢) ، (٨،٣) ، (٨،٤) على الترتيب أمثلة على  $r = +1$  و  $r = -1$  تساوى تقريبا صفراً .



من الواضح أنه يمكن استخدام الكمية  $b/\sqrt{S_{xx}}/\sqrt{S_{yy}}$  والتي سنرمز لها من الآن فصاعدا بالرمز  $r$  كتقدير لمعامل الارتباط  $P$ .

### تعريف معامل الارتباط The correlation coefficient

إن القياس  $r$  لدرجة العلاقة الخطية بين متغيرين  $X$  و  $Y$  تقدر بواسطة معامل الارتباط العيني  $r$  حيث :

$$r = b \sqrt{\frac{S_{xx}}{S_{yy}}} = \frac{S_{xy}}{\sqrt{S_{xx} S_{yy}}}$$

وتنحصر قيمة  $r$  بين  $-1$  ،  $+1$  ، ويجب أن نكون حريصين في تفسيراتنا ، فمثلا قيم  $r$  المساوية لـ  $0.2$  و  $0.4$  تعني فقط أن لدينا إرتباطين موجبين وأحدهما أقوى من الآخر إلى حد ما ، ومن الخطأ أن نستنتج أن  $r = 0.4$  تشير إلى وجود علاقة خطية أفضل بمرتين من التي تشير إليها القيمة  $r = 0.2$  ، ومن ناحية أخرى إذا افترضنا  $r^2$  فإن  $r^2 = 100\%$  من الاختلافات في قيم  $Y$  يمكن أن تحسب بواسطة العلاقة الخطية بالمتغير  $X$  . وهكذا فإن الارتباط  $0.4$  يعني أن  $16\%$  من الاختلافات في المتغير  $Y$  تحسب بواسطة الفروق في المتغير  $X$  .

مثال (٨,١٠)

إذا علمت أن أطوال ثمانية آباء وأطوال أكبر أولادهم معطاة بالجدول (٨,٥) :

جدول (٨,٥)

طول الأب بالانثى $x$	طول أكبر أولاده $y$
63	65
64	67
70	69
72	70
65	64
67	68
68	71
66	63

- أ — أوجد من أجل هذه المعلومات خط الانحدار  $y$  على  $x$  .  
 ب — أوجد تقديراً لطول أكبر الأولاد وذلك من أجل طول الأب 69 انش .  
 ج — ارسم خط الانحدار .  
 د — احسب معامل الارتباط العيني  $r$  ثم فسر هذا الارتباط .

الحل

أولاً : من أجل إيجاد خط الانحدار نلاحظ أن :

الجدول (٨,٦)

$x$	$y$	$xy$	$x^2$	$y^2$	$\bar{y}$	$y - \bar{y}$	$(y - \bar{y})^2$
63	65	4095	3969	4225	64.55	0.45	0.2025
64	67	4288	4096	4489	65.21	1.79	3.2041
70	69	4830	4900	4761	69.20	-0.20	0.0400
72	70	5040	5184	4900	70.53	-0.53	0.2809
65	64	4160	4225	4096	65.88	-1.88	3.5344
67	68	4556	4489	4624	67.21	.78	0.6241
68	71	4828	4624	5041	76.87	3.13	9.7969
66	63	4158	4356	3969	66.54	-3.54	12.5316
535	537	35955	35843	36105			30.2145

كما نلاحظ أن  $\bar{y} = \frac{537}{8} = 67.125$  ,  $\bar{x} = \frac{535}{8} = 66.875$  ومن المعلوم أن معادلة انحدار  $Y$  على  $X$  هي :

$$\hat{y} = a + bx$$

كما نلاحظ أن :

$$b = \frac{\sum_{i=1}^8 x_i y_i - n \bar{x} \bar{y}}{\sum_{i=1}^8 x_i^2 - n \bar{x}^2} = \frac{35955 - 8(66.875)(67.125)}{35843 - 8(66.875)^2}$$

وأن :

$$a = \bar{y} - b\bar{x} = 67.125 - (0.6647)(66.875) = 22.6732$$

لذلك فإن معادلة الانحدار المطلوبة هي :

$$\hat{y} = 22.6732 + 0.6647x$$

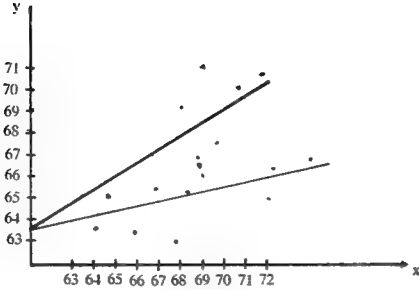
ثانيا : باستخدام معادلة الانحدار من أجل  $x = 69$  ، فإننا نجد أن :

$$\hat{y} = 22.6732 + 0.6647(69) = 68.5373$$

ثالثا : من أجل رسم الانحدار السابق ، فإننا نحتاج إلى نقطتين فقط ، نلاحظ مثلاً أنه من أجل :

$$\begin{array}{ll} \hat{y} = 63.8846 & \text{أن} \quad x = 62 \\ \hat{y} = 68.5375 & \text{أن} \quad x = 69 \end{array}$$

وسنحصل عندئذ على خط الانحدار  $\hat{y} = 22.6732 + 0.6647x$  بالوصل بين النقطتين  $N(69, 68.5375)$  ,  $M(62, 63.8846)$  . يوضح الشكل (٨,٥) مخطط الانتشار وخط الانحدار المطلوب .



الشكل (٨، ٥)

نلاحظ أيضاً أن :

$$S_{xx} = \sum_{i=1}^8 x_i^2 - \frac{(\sum_{i=1}^8 x_i)^2}{8} = 64.875$$

$$S_{yy} = \sum_{i=1}^8 y_i^2 - \frac{(\sum_{i=1}^8 y_i)^2}{8} = 58.875$$

$$S_{xy} = \sum_{i=1}^8 x_i y_i - \frac{(\sum_{i=1}^8 x_i)(\sum_{i=1}^8 y_i)}{8} = 43.125$$

لذلك فإن :

$$r = \frac{S_{xy}}{\sqrt{S_{xx} S_{yy}}} = \frac{43.125}{\sqrt{(64.875)(58.875)}}$$

ومنه :

$$r = 6977$$

يشير معامل الارتباط السابق إلى وجود علاقة خطية بين X و Y ، وحيث إن  $r^2 = 0.4869$  ، لذلك فإن بإمكاننا القول بأن 48.69% تقريبا من الاختلافات في قيم y تحسب بواسطة علاقة انحدار Y على X .

لاختبار الفرضية  $H_0: \rho = \rho_0$  ضد الفرضية البديلة  $H_1: \rho \neq \rho_0$  ، فإننا نلاحظ أنه من ملاحظات التوزيع الطبيعي ذى البعدين فإن الكمية  $\frac{1}{2} \ln \frac{1+r}{1-r}$  تمثل قيمة متغير عشوائى يتبع تقريبا التوزيع الطبيعي بالوسط  $\frac{1}{2} \ln \frac{1+\rho_0}{1-\rho_0}$  والتباين  $\frac{1}{n-3}$  ، وهكذا فإن إجراء الاختبار هو فى حساب :

$$z = \frac{\sqrt{n-3}}{2} \left[ \ln \left( \frac{1+r}{1-r} \right) - \ln \left( \frac{1+\rho_0}{1-\rho_0} \right) \right]$$

$$= \frac{\sqrt{n-3}}{n} \ln \left[ \frac{(1+r)(1-\rho_0)}{(1-r)(1+\rho_0)} \right]$$

ومقارنته مع النقاط الحرجة للتوزيع الطبيعي المعيارى .

مثال (٨، ١١)

اختبر فى المثال السابق الفرضية فى أنه ليس هناك ارتباط خطى بين المتغيرين .  
استخدم لذلك مستوى معنوية  $\alpha = 0.05$  .

الحل

- |                      |                          |
|----------------------|--------------------------|
| $H_0: \rho = 0$      | ١ — إن فرض العدم هو      |
| $H_1: \rho \neq 0$   | ٢ — إن الفرض البديل هو   |
| $\alpha = 0.05$      | ٣ —                      |
| $Z < 1.96, Z > 1.96$ | ٤ — إن المنطقة الحرجة هي |
|                      | ٥ — الحسابات             |

$$z = \frac{\sqrt{5}}{2} \ln \left[ \frac{1.6978}{0.3022} \right] = 1.9297$$

٦ — الاستنتاج : إننا نقبل الفرضية بعدم وجود علاقة خطية .

## المراجع

- Harr, D.R. and Zehna, P.W.** (1971) *Probability*, California: Brooks/Cole Publishing Co
- Bowker, A.H. and Lieberman, G.J.** (1972) *Engineering Statistics*, 3rd ed., Englewood Cliffs, Prentice Hall Inc
- Feller, William** (1966) *An Introduction to Probability and its Applications*, New York: John Wiley, Vol. 1, 1968, 3rd ed., and Vol. 2.
- Fisz, Murek** (1963) *Probability Theory and Mathematical Statistics*, 3rd ed., New York: John Wiley.
- Freund, J.E.** (1971) *Mathematical Statistics*, 2nd ed., Englewood Cliffs, Prentice-Hall Inc.
- Guttman, I.S.S. Wilks and Hunter, J.S.** (1971) *Introductory Engineering Statistics*, 2nd ed., New York: John Wiley and Sons.
- Hawkins, C.A. and Weber, J.E.** (1980) *Statistical Analysis*, New York: Harper and Row.
- Larson, Harold J.** (1973) *Introduction to the Theory of Statistics*, New York: John Wiley.
- McClave, J.T. and Dietrich, F.H.** (1979) *Statistics*, San Francisco, Dellen Publishing Co.
- Mendenhall, W., Scheaffer, R.L. and Wackerly, D.D.** (1981) *Mathematical Statistics with Applications*, 2nd ed., Boston: Duxbury Press.
- Miller, I. and Freund, J.E.** (1977). *Probability and Statistics for Engineers*, 2nd ed., Englewood Cliffs, Prentice-Hall, Inc.
- Rao, C.R.** (1973) *Linear Statistical Inference and its Applications*, 2nd ed., New York: John Wiley.
- Scheaffer, Richard, L. and McClave, James T.** (1982) *Statistics for Engineers*, Boston: Duxbury Press.
- Stephens, M.A.** (1974) EDF statistics for goodness and fit and some comparisons. *Journal of Am. Sta Assn* . Vol. 69, no. 347, pp. 730-737
- Walpole, R.E. and Myers, R.H.** (1978) *Probability and Statistics for Engineers and Scientists*, 2nd ed., New York, Macmillan Publishing Co.
- Zehna, Peter W.** (1970) *Probability Distributions and Statistics*. Boston: Allyn and Bacon, Inc.



# المشلاحق

- جدول I : المربعات والجذور التربيعية ■ جدول II : مجموع الاحتمال الحدائي
- $\sum_{x=0}^r b(x;n,p)$  ■ جدول III : مجموع الاحتمال البواسوني  $b(x;n)$
- $\sum_{x=0}^r$  ■ جدول IV : المساحة تحت المنحنى الطبيعي ■ جدول V :
- القيم الحرجة في توزيع ■ جدول VI : القيم الحرجة في توزيع كاي مربع
- جدول VII : القيم الحرجة في توزيع t ■ جدول VIII : عوامل التحميل في التوزيع الطبيعي .



جدول ١ : المربعات والجذور التربيعية

$n$	$n^2$	$\sqrt{n}$	$\sqrt{10n}$	$n$	$n^2$	$\sqrt{n}$	$\sqrt{10n}$
1.0	1.00	1.000	3.162	5.5	30.25	2.345	7.416
1.1	1.21	1.049	3.317	5.6	31.36	2.366	7.483
1.2	1.44	1.095	3.464	5.7	32.49	2.387	7.550
1.3	1.69	1.140	3.606	5.8	33.64	2.408	7.616
1.4	1.96	1.183	3.742	5.9	34.81	2.429	7.681
1.5	2.25	1.225	3.873	6.0	36.00	2.449	7.746
1.6	2.56	1.265	4.000	6.1	37.21	2.470	7.810
1.7	2.89	1.304	4.123	6.2	38.44	2.490	7.874
1.8	3.24	1.342	4.243	6.3	39.69	2.510	7.937
1.9	3.61	1.378	4.359	6.4	40.96	2.530	8.000
2.0	4.00	1.414	4.472	6.5	42.25	2.550	8.062
2.1	4.41	1.449	4.583	6.6	43.56	2.569	8.124
2.2	4.84	1.483	4.690	6.7	44.89	2.588	8.185
2.3	5.29	1.517	4.796	6.8	46.24	2.608	8.246
2.4	5.76	1.549	4.899	6.9	47.61	2.627	8.307
2.5	6.25	1.581	5.000	7.0	49.00	2.646	8.367
2.6	6.76	1.612	5.099	7.1	50.41	2.66	8.426
2.7	7.29	1.643	5.196	7.2	51.84	2.68	8.485
2.8	7.84	1.673	5.292	7.3	53.29	2.702	8.544
2.9	8.41	1.703	5.385	7.4	54.76	2.720	8.602
3.0	9.00	1.732	5.477	7.5	56.25	2.739	8.660
3.1	9.61	1.761	5.568	7.6	57.76	2.757	8.718
3.2	10.24	1.789	5.657	7.7	59.29	2.775	8.775
3.3	10.89	1.817	5.745	7.8	60.84	2.793	8.832
3.4	11.56	1.844	5.831	7.9	62.41	2.811	8.888
3.5	12.25	1.871	5.916	8.0	64.00	2.828	8.944
3.6	12.96	1.897	6.000	8.1	65.61	2.846	9.000
3.7	13.69	1.924	6.083	8.2	67.24	2.864	9.055
3.8	14.44	1.949	6.164	8.3	68.89	2.881	9.110
3.9	15.21	1.975	6.245	8.4	70.56	2.898	9.165
4.0	16.00	2.000	6.325	8.5	72.25	2.915	9.220
4.1	16.81	2.025	6.403	8.6	73.96	2.933	9.274
4.2	17.64	2.049	6.481	8.7	75.69	2.950	9.327
4.3	18.49	2.074	6.557	8.8	77.44	2.966	9.381
4.4	19.36	2.098	6.633	8.9	79.21	2.983	9.434
4.5	20.25	2.121	6.708	9.0	81.00	3.000	9.487
4.6	21.16	2.145	6.782	9.1	82.81	3.017	9.539
4.7	22.09	2.168	6.856	9.2	84.64	3.033	9.592
4.8	23.04	2.191	6.928	9.3	86.49	3.050	9.644
4.9	24.01	2.214	7.000	9.4	88.36	3.066	9.695
5.0	25.00	2.236	7.071	9.5	90.25	3.082	9.747
5.1	26.01	2.258	7.141	9.6	92.16	3.098	9.798
5.2	27.04	2.280	7.211	9.7	94.09	3.114	9.849
5.3	28.09	2.302	7.280	9.8	96.04	3.130	9.899
5.4	29.16	2.324	7.348	9.9	98.01	3.146	9.950

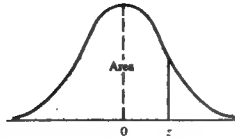
جدول ١١ : مجموع الاحتمال الخداني

		P									
n	p	0.10	0.20	0.25	0.30	0.40	0.50	0.60	0.70	0.80	0.90
5	0	0.5905	0.3277	0.2373	0.1681	0.0778	0.0312	0.0102	0.0024	0.0003	0.0000
	1	0.9185	0.7373	0.6328	0.5282	0.3750	0.1875	0.0870	0.0308	0.0067	0.0000
	2	0.9914	0.9431	0.8995	0.8369	0.6825	0.4174	0.1906	0.0879	0.0300	0.0000
	3	1.0000	0.9997	0.9990	0.9976	0.9988	0.9688	0.9222	0.8319	0.6723	0.4615
	4	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
10	0	0.3487	0.1074	0.0563	0.0282	0.0060	0.0010	0.0001	0.0000	0.0000	0.0000
	1	0.7361	0.3758	0.2440	0.1493	0.0464	0.0107	0.0017	0.0001	0.0000	0.0000
	2	0.9598	0.7678	0.6256	0.4828	0.3047	0.0123	0.0016	0.0001	0.0000	0.0000
	3	0.9913	0.9791	0.9749	0.9685	0.9585	0.9379	0.9106	0.8679	0.7884	0.6667
	4	1.0000	0.9972	0.9921	0.9847	0.9631	0.9370	0.1662	0.0474	0.0064	0.0002
	5	0.9999	0.9936	0.9803	0.9527	0.8338	0.6230	0.3669	0.1503	0.0328	0.0016
	6	1.0000	0.9991	0.9965	0.9874	0.9452	0.8284	0.6717	0.3504	0.1209	0.0128
	7	1.0000	0.9999	0.9996	0.9994	0.9977	0.9843	0.9513	0.8723	0.7532	0.6011
	8	1.0000	1.0000	1.0000	0.9999	0.9983	0.9893	0.9536	0.8507	0.6242	0.2639
	9	1.0000	1.0000	1.0000	1.0000	0.9999	0.9990	0.9940	0.9718	0.8526	0.6513
	10	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
15	0	0.2059	0.0352	0.0134	0.0047	0.0005	0.0000	0.0000	0.0000	0.0000	0.0000
	1	0.5490	0.1671	0.0802	0.0353	0.0052	0.0005	0.0000	0.0000	0.0000	0.0000
	2	0.8159	0.3980	0.2361	0.1268	0.0271	0.0037	0.0003	0.0000	0.0000	0.0000
	3	0.9444	0.4482	0.4613	0.3769	0.2903	0.0176	0.0019	0.0001	0.0000	0.0000
	4	0.9873	0.8358	0.6865	0.5155	0.2173	0.0592	0.0094	0.0007	0.0000	0.0000
	5	0.9978	0.9389	0.8516	0.7216	0.4632	0.1509	0.0338	0.0037	0.0001	0.0000
	6	0.9997	0.9819	0.9434	0.8689	0.6096	0.3036	0.0951	0.0152	0.0008	0.0000
	7	1.0000	0.9958	0.9827	0.9500	0.7869	0.5000	0.2131	0.0500	0.0042	0.0000
	8	1.0000	0.9992	0.9958	0.9848	0.9050	0.6964	0.3902	0.1331	0.0181	0.0023
	9	1.0000	0.9999	0.9992	0.9963	0.9662	0.8491	0.6968	0.2784	0.0511	0.0023
	10	1.0000	1.0000	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
	11	1.0000	1.0000	1.0000	1.0000	0.9999	0.9981	0.9824	0.9051	0.7518	0.0556
	12	1.0000	1.0000	1.0000	1.0000	0.9997	0.9963	0.9729	0.8732	0.6020	0.1841
	13	1.0000	1.0000	1.0000	1.0000	0.9999	0.9995	0.9948	0.9647	0.8329	0.4510
	14	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9998	0.9979	0.9948	0.9900
	15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
20	0	0.1216	0.0115	0.0032	0.0006	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	1	0.3917	0.0692	0.0241	0.0087	0.0005	0.0000	0.0000	0.0000	0.0000	0.0000
	2	0.6769	0.2061	0.0913	0.0355	0.0036	0.0002	0.0000	0.0000	0.0000	0.0000
	3	0.8570	0.4114	0.2252	0.1071	0.0160	0.0013	0.0001	0.0000	0.0000	0.0000
	4	0.9364	0.6296	0.4148	0.2375	0.0510	0.0059	0.0003	0.0000	0.0000	0.0000
	5	0.9617	0.7046	0.4617	0.2466	0.0616	0.0080	0.0016	0.0000	0.0000	0.0000
	6	0.9776	0.9133	0.7858	0.6080	0.2500	0.0577	0.0063	0.0003	0.0000	0.0000
	7	0.9996	0.9679	0.8982	0.7723	0.4159	0.1316	0.0210	0.0013	0.0000	0.0000
	8	0.9999	0.9990	0.9991	0.9867	0.9596	0.2517	0.0565	0.0051	0.0001	0.0000
	9	1.0000	0.9997	0.9986	0.9952	0.9729	0.7127	0.4173	0.1627	0.0421	0.0000
	10	1.0000	0.9994	0.9961	0.9829	0.8725	0.5881	0.2447	0.0480	0.0026	0.0000
	11	1.0000	0.9999	0.9991	0.9949	0.9435	0.7683	0.4044	0.1133	0.0100	0.0001
	12	1.0000	1.0000	0.9998	0.9987	0.9790	0.8468	0.5841	0.2277	0.0321	0.0004
	13	1.0000	1.0000	1.0000	0.9991	0.9953	0.9423	0.8000	0.5067	0.1867	0.0300
	14	1.0000	1.0000	1.0000	1.0000	0.9984	0.9791	0.8494	0.5836	0.1598	0.0412
	15	1.0000	1.0000	1.0000	1.0000	0.9997	0.9941	0.9490	0.7625	0.3704	0.0833
	16	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9940	0.9429	0.5886	0.1330
	17	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998	0.9964	0.9423	0.5839	0.1339
	18	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9995	0.9924	0.9398	0.6083
	19	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9992	0.9885	0.8784
	20	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000









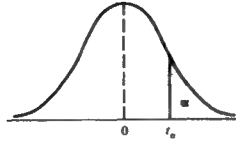
جدول IV : المساحة تحت المنحنى الطبيعي.

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002
3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003
3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005
3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007
3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
2.9	0.0019	0.0018	0.0017	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
1.8	0.0359	0.0352	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0722	0.0708	0.0694	0.0681
1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
0.6	0.2743	0.2709	0.2676	0.2643	0.2610	0.2578	0.2546	0.2514	0.2483	0.2451
0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8769	0.8789	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9278	0.9292	0.9306	0.9319



جدول IV: المساحة تحت المنحنى الطبيعي (تابع)

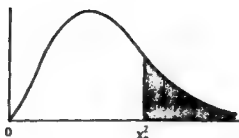
[illegible]



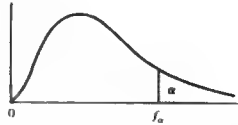
جدول ٧ : القيم الحرجة في توزيع t

$\nu$	$\alpha$				
	0.10	0.05	0.025	0.01	0.005
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.179	2.681	3.055
13	1.350	1.771	2.160	2.650	3.012
14	1.345	1.761	2.145	2.624	2.977
15	1.341	1.753	2.131	2.602	2.947
16	1.337	1.746	2.120	2.583	2.921
17	1.333	1.740	2.110	2.567	2.898
18	1.330	1.734	2.101	2.552	2.878
19	1.328	1.729	2.093	2.539	2.861
20	1.325	1.725	2.086	2.528	2.845
21	1.323	1.721	2.080	2.518	2.831
22	1.321	1.717	2.074	2.508	2.819
23	1.319	1.714	2.069	2.500	2.807
24	1.318	1.711	2.064	2.492	2.797
25	1.316	1.708	2.060	2.485	2.787
26	1.315	1.706	2.056	2.479	2.779
27	1.314	1.703	2.052	2.473	2.771
28	1.313	1.701	2.048	2.467	2.763
29	1.311	1.699	2.045	2.462	2.756
inf.	1.282	1.645	1.960	2.326	2.576

جدول VI : القيم الحرجة في توزيع كاي مربع



ν	α							
	0.995	0.99	0.975	0.95	0.05	0.025	0.01	0.005
1	0.00393	0.00157	0.00982	0.00393	3.841	5.024	6.635	7.879
2	0.0100	0.0201	0.0506	0.103	5.991	7.378	9.210	10.597
3	0.0717	0.115	0.216	0.352	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	11.070	12.832	15.086	16.750
6	0.676	0.872	1.237	1.635	12.592	14.449	16.812	18.548
7	0.989	1.239	1.690	2.157	14.067	16.013	18.475	20.278
8	1.344	1.646	2.180	2.733	15.507	17.535	20.090	21.955
9	1.735	2.088	2.700	3.325	16.919	19.023	21.666	23.589
10	2.156	2.558	3.247	3.940	18.307	20.483	23.209	25.188
11	2.603	3.053	3.816	4.575	19.675	21.920	24.725	26.757
12	3.074	3.571	4.404	5.226	21.026	23.337	26.217	28.300
13	3.565	4.107	5.009	5.892	22.362	24.736	27.688	29.819
14	4.075	4.660	5.629	6.571	23.685	26.119	29.141	31.319
15	4.601	5.229	6.262	7.261	24.996	27.488	30.578	32.801
16	5.142	5.812	6.908	7.962	26.296	28.845	32.000	34.267
17	5.697	6.408	7.564	8.672	27.587	30.191	33.409	35.718
18	6.265	7.015	8.231	9.390	28.869	31.526	34.805	37.156
19	6.844	7.633	8.907	10.117	30.144	32.852	36.191	38.582
20	7.434	8.260	9.591	10.851	31.410	34.170	37.566	39.997
21	8.034	8.897	10.283	11.591	32.671	35.479	38.932	41.401
22	8.643	9.542	10.982	12.338	33.924	36.781	40.289	42.796
23	9.260	10.196	11.689	13.091	35.172	38.076	41.638	44.181
24	9.886	10.856	12.401	13.848	36.415	39.364	42.980	45.558
25	10.520	11.524	13.120	14.611	37.652	40.646	44.314	46.928
26	11.160	12.196	13.844	15.379	38.885	41.923	45.642	48.290
27	11.808	12.879	14.573	16.151	40.113	43.194	46.963	49.645
28	12.461	13.565	15.308	16.928	41.337	44.461	48.278	50.993
29	13.121	14.256	16.047	17.708	42.557	45.722	49.588	52.336
30	13.787	14.953	16.791	18.493	43.773	46.979	50.892	53.672

جداول VII : القيم الحرجة في توزيع  $\chi^2$  $f_{0.05}(\nu_1, \nu_2)$ 

$\nu_2$	$\nu_1$								
	1	2	3	4	5	6	7	8	9
1	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04
120	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96
$\infty$	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88

جدول VII : القيم الحرجة في توزيع  $f$  (تابع)

$$f_{0.05}(v_1, v_2)$$

$v_2$	$v_1$									
	10	12	15	20	24	30	40	60	120	$\infty$
1	241.9	243.9	245.9	248.0	249.1	250.1	251.1	252.2	253.3	254.3
2	19.40	19.41	19.43	19.45	19.45	19.46	19.47	19.48	19.49	19.50
3	8.79	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55	8.53
4	5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.63
5	4.74	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.36
6	4.06	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70	3.67
7	3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27	3.24
8	3.35	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97	2.94
9	3.14	3.07	3.01	2.94	2.90	2.86	2.83	2.79	2.75	2.71
10	2.98	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.58	2.54
11	2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.40
12	2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	2.30
13	2.67	2.60	2.53	2.46	2.42	2.38	2.34	2.30	2.25	2.21
14	2.60	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	2.13
15	2.54	2.48	2.40	2.33	2.29	2.25	2.20	2.16	2.11	2.07
16	2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.01
17	2.45	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.01	1.96
18	2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.92
19	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88
20	2.35	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90	1.84
21	2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.92	1.87	1.81
22	2.30	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	1.78
23	2.27	2.20	2.13	2.05	2.01	1.96	1.91	1.86	1.81	1.76
24	2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79	1.73
25	2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77	1.71
26	2.22	2.15	2.07	1.99	1.95	1.90	1.85	1.80	1.75	1.69
27	2.20	2.13	2.06	1.97	1.93	1.88	1.84	1.79	1.73	1.67
28	2.19	2.12	2.04	1.96	1.91	1.87	1.82	1.77	1.71	1.65
29	2.18	2.10	2.03	1.94	1.90	1.85	1.81	1.75	1.70	1.64
30	2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.68	1.62
40	2.08	2.00	1.92	1.84	1.79	1.74	1.69	1.64	1.58	1.51
60	1.99	1.92	1.84	1.75	1.70	1.65	1.59	1.53	1.47	1.39
120	1.91	1.83	1.75	1.66	1.61	1.55	1.50	1.43	1.35	1.25
$\infty$	1.83	1.75	1.67	1.57	1.52	1.46	1.39	1.32	1.22	1.00

جداول VII : القيم الحرجة في توزيع  $f$  (تابع)

$f_{\alpha, n_1, n_2}$

$n_2$	$n_1$								
	1	2	3	4	5	6	7	8	9
1	4052	4999.5	5403	5625	5764	5859	5928	5981	6022
2	98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39
3	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.35
4	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66
5	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16
6	13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98
7	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72
8	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91
9	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35
10	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39
13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19
14	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78
17	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68
18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60
19	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46
21	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40
22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35
23	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30
24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26
25	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.22
26	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18
27	7.68	5.49	4.60	4.11	3.78	3.56	3.39	3.26	3.15
28	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12
29	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.09
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72
120	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56
$\infty$	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41

جدول VII : القيم الحرجة في توزيع  $f$  (تابع)

$$f_{0.01}(v_1, v_2)$$

$v_2$	$v_1$									
	10	12	15	20	24	30	40	60	120	$\infty$
1	6056	6106	6157	6209	6235	6261	6287	6313	6339	6366
2	99.40	99.42	99.43	99.45	99.46	99.47	99.47	99.48	99.49	99.50
3	27.23	27.05	26.87	26.69	26.60	26.50	26.41	26.32	26.22	26.13
4	14.55	14.37	14.20	14.02	13.93	13.84	13.75	13.65	13.56	13.46
5	10.05	9.89	9.72	9.55	9.47	9.38	9.29	9.20	9.11	9.02
6	7.87	7.72	7.56	7.40	7.31	7.23	7.14	7.06	6.97	6.88
7	6.62	6.47	6.31	6.16	6.07	5.99	5.91	5.82	5.74	5.65
8	5.81	5.67	5.52	5.36	5.28	5.20	5.12	5.03	4.95	4.86
9	5.26	5.11	4.96	4.81	4.73	4.65	4.57	4.48	4.40	4.31
10	4.85	4.71	4.56	4.41	4.33	4.25	4.17	4.08	4.00	3.91
11	4.54	4.40	4.25	4.10	4.02	3.94	3.86	3.78	3.69	3.60
12	4.30	4.16	4.01	3.86	3.78	3.70	3.62	3.54	3.45	3.36
13	4.10	3.96	3.82	3.66	3.59	3.51	3.43	3.34	3.25	3.17
14	3.94	3.80	3.66	3.51	3.43	3.35	3.27	3.18	3.09	3.00
15	3.80	3.67	3.52	3.37	3.29	3.21	3.13	3.05	2.96	2.87
16	3.69	3.55	3.41	3.26	3.18	3.10	3.02	2.93	2.84	2.75
17	3.59	3.46	3.31	3.16	3.08	3.00	2.92	2.83	2.75	2.65
18	3.51	3.37	3.23	3.08	3.00	2.92	2.84	2.75	2.66	2.57
19	3.43	3.30	3.15	3.00	2.92	2.84	2.76	2.67	2.58	2.49
20	3.37	3.23	3.09	2.94	2.86	2.78	2.69	2.61	2.52	2.42
21	3.31	3.17	3.03	2.88	2.80	2.72	2.64	2.55	2.46	2.36
22	3.26	3.12	2.98	2.83	2.75	2.67	2.58	2.50	2.40	2.31
23	3.21	3.07	2.93	2.78	2.70	2.62	2.54	2.45	2.35	2.26
24	3.17	3.03	2.89	2.74	2.66	2.58	2.49	2.40	2.31	2.21
25	3.13	2.99	2.85	2.70	2.62	2.54	2.45	2.36	2.27	2.17
26	3.09	2.96	2.81	2.66	2.58	2.50	2.42	2.33	2.23	2.13
27	3.06	2.93	2.78	2.63	2.55	2.47	2.38	2.29	2.20	2.10
28	3.03	2.90	2.75	2.60	2.52	2.44	2.35	2.26	2.17	2.06
29	3.00	2.87	2.73	2.57	2.49	2.41	2.33	2.23	2.14	2.03
30	2.98	2.84	2.70	2.55	2.47	2.39	2.30	2.21	2.11	2.01
40	2.80	2.66	2.52	2.37	2.29	2.20	2.11	2.02	1.92	1.80
60	2.63	2.50	2.35	2.20	2.12	2.03	1.94	1.84	1.73	1.60
120	2.47	2.34	2.19	2.03	1.95	1.86	1.76	1.66	1.53	1.38
$\infty$	2.32	2.18	2.04	1.88	1.79	1.70	1.59	1.47	1.32	1.00

## جدول VIII : عوامل العمل في التوزيع الطبيعي

 $v = 0.95$ 

n	1 - $\alpha$		
	0.90	0.95	0.99
2	32.019	37.674	48.430
3	8.380	9.916	12.861
4	5.369	6.370	8.299
5	4.275	5.079	6.634
6	3.712	4.414	5.775
7	3.369	4.007	5.248
8	3.136	3.732	4.891
9	2.967	3.532	4.631
10	2.839	3.379	4.433
11	2.737	3.259	4.277
12	2.655	3.162	4.150
13	2.587	3.081	4.044
14	2.529	3.012	3.955
15	2.480	2.954	3.878
16	2.437	2.903	3.812
17	2.400	2.858	3.754
18	2.366	2.819	3.702
19	2.337	2.784	3.656
20	2.310	2.752	3.615
25	2.208	2.631	3.457
30	2.140	2.549	3.350
35	2.090	2.490	3.272
40	2.052	2.445	3.213

 $v = 0.99$ 

n	1 - $\alpha$		
	0.90	0.95	0.99
2	160.193	188.491	242.300
3	18.930	22.401	29.055
4	9.398	11.150	14.527
5	6.612	7.855	10.260
6	5.337	6.345	8.301
7	4.613	5.488	7.187
8	4.147	4.936	6.468
9	3.822	4.550	5.966
10	3.582	4.265	5.594
11	3.397	4.045	5.308
12	3.250	3.870	5.079
13	3.130	3.727	4.893
14	3.029	3.608	4.737
15	2.945	3.507	4.605
16	2.872	3.421	4.492
17	2.808	3.345	4.393
18	2.753	3.279	4.307
19	2.703	3.221	4.230
20	2.659	3.168	4.161
25	2.494	2.972	3.904
30	2.385	2.841	3.733
35	2.306	2.748	3.611
40	2.247	2.677	3.518



## جدول VIII : عوامل التحميل في التوزيع الطبيعي ( تابع )

 $v = 0.95$ 

n	1 - $\alpha$		
	0.90	0.95	0.99
45	2.021	2.408	3.165
50	1.996	2.379	3.126
55	1.976	2.354	3.094
60	1.958	2.333	3.066
65	1.943	2.315	3.042
70	1.929	2.299	3.021
75	1.917	2.285	3.002
80	1.907	2.272	2.986
85	1.897	2.261	2.971
90	1.889	2.251	2.958
95	1.881	2.241	2.945
100	1.874	2.233	2.934
150	1.825	2.175	2.859
200	1.798	2.143	2.816
250	1.780	2.121	2.788
300	1.767	2.106	2.767
400	1.749	2.084	2.739
500	1.737	2.070	2.721
600	1.729	2.060	2.707
700	1.722	2.052	2.697
800	1.717	2.046	2.688
900	1.712	2.040	2.682
1000	1.709	2.036	2.676
$\infty$	1.645	1.960	2.576

 $v = 0.99$ 

n	1 - $\alpha$		
	0.90	0.95	0.99
45	2.200	2.621	3.444
50	2.162	2.576	3.385
55	2.130	2.538	3.335
60	2.103	2.506	3.293
65	2.080	2.478	3.257
70	2.060	2.454	3.225
75	2.042	2.433	3.197
80	2.026	2.414	3.173
85	2.012	2.397	3.150
90	1.999	2.382	3.130
95	1.987	2.368	3.112
100	1.977	2.355	3.096
150	1.905	2.270	2.983
200	1.865	2.222	2.921
250	1.839	2.191	2.880
300	1.820	2.169	2.850
400	1.794	2.138	2.809
500	1.777	2.117	2.783
600	1.764	2.102	2.763
700	1.755	2.091	2.748
800	1.747	2.082	2.736
900	1.741	2.075	2.726
1000	1.736	2.068	2.718
$\infty$	1.645	1.960	2.576



# ثبت المصطلحات

■ عربي / إنجليزي ■ إنجليزي / عربي .



## عربي/الانجليزي

Estimation	تقدير	١	
Unbiased estimator	تقدير غير متحيز	Probability	احتمال
Biased estimator	تقدير متحيز	Conditional probability	احتمال شرطي
Approximation	تقريب	Statistics	إحصاء
Repetition	تكرار	Deductive statistics	الإحصاء الاستنتاجي
Relative frequency	التكرار النسبي	Descriptive statistics	إحصاء وصفي
Probability distribution	توزيع احتمالي	Two-sided test	اختبار ثنائي الجانب
Continuous probability	توزيع احتمالي مستمر	Two-tailed test	اختبار ثنائي الذيل
Joint probability distribution	توزيع احتمالي مشترك	One-sided test	اختبار وحيد الجانب
Bernoulli distribution	توزيع برنولي	One-tailed test	اختبار وحيد الذيل
Empericall distribution	توزيع تجريبي	Variation	الاختلاف
Cumulative distribution	توزيع تراكمي	Selection	اختيار
Negative distribution	توزيع سالب	Illness	أساس
Conditional distribution	توزيع شرطي	Dispersion	الانتشار
Sampling	التوزيع العيني للوسط	Bending	انحناء
distribution of mean		Standard deviation	الانحراف المعياري
Sampling	التوزيع العيني للنسبة		
distribution of proportion			
Normal distribution	توزيع طبيعي	ت	
Gamma distribution	توزيع غاما	Permutations	تباديل
Gaussian distribution	توزيع غوس	Variance	تباين
Chi-Square distribution	توزيع كاي - مربع	Partition	تجزئة
Sampling	توزيع المعاينة	Orderd partition	تجزيات مرتبة
Exprimental sampling	توزيع المعاينة التجريبي	Analysis	تحليل
Marginal distribution	توزيع هامشي	Regression	الترجع
Hypergeometric	توزيع هندسي زائدي	Counting	الترقيم - العدد
Expectation	التوقع	Intersection	تقاطع

Angle	ز	ثابت	Constant	ث
		زاوية		
Negative	ص	جلاء (ضرب)	Product	ج
Angular velocity		السرعة الزاوية		
Sampling without replacement		السحب بدون الإعادة		ح
Sampling with replacement		السحب مع الإعادة	Event	حادث
Chain		سلسلة	Elementary event	حادث أولي
			Impossible event	حادث مستحيل
	ف		Binomial	حتماني
Vector		شعاع	Mutually exclusive events	حوادث متنافية تبادلياً
Graph		شكل بياني	Independent events	حوادث مستقلة
Object		شيء - هدف		
	ض		Asymptote	خط تنازلي
Product		(ضرب - جلاء)		ا
			Probability function	دالة احتمالية
	ع		Frequency function	دالة تكرارية
Counting		المد - الترقيم	Algebraic function	دالة جبرية
Moment		عزم	Linear function	دالة خطية
Moment about zero		عزم حول الصفر	Step function	دالة درجية
Moment about mean		عزم حول وسط	Density function	دالة الكثافة
Column		عمود	Probability density function	دالة الكثافة الاحتمالية
Member		عضو	Continuous function	دالة مستمرة
Component		عنصر	Degrees of freedom	درجات من الحرية
Sample		عينة		ذ
Random sample		عينة عشوائية	Tail	الذيل ( وجه الكتابة في قطعة النقود )
	غ			ر
Uncountable		غير قابل للمد	Head	الرأس ( وجه النقش في قطعة النقود )
Infinite		غير محدود (لانهاي)		

Finite	محدود	ف	
Transformed	محول	Alternative hypothesis	الفرض البديل
Venn diagram	مخطط فين	Null hypothesis	فرض العدم
Independent	مستقل	Hypotheses	فرضيات
Continuous	مستمر	Statistical hypotheses	الفرضيات الاحصائية
Matrix	مصفوفة		
Defective	معيب	Sample space	فضاء العينة
Observation	ملاحظة	Binomial expansion	فك حيداني
Volume relatives	متناسب الحجم		
Critical region	منطقة حرجية	ق	
Region of rejection of hypothesis	منطقة رفض الفرضية	Countable	قابل للعد
		Statistical decision	القرارات الاحصائية
		Power	قوة
		Critical number	قيمة حرجية
		Expected value	قيمة متوقعة
Decisions Theory	نظرية القرارات		
		م	
Analytic Geometry	هندسة تحليلية	Inequality	متباينة
		Sequence	متتالية
		Independent variables	متغيرات مستقلة
		Random variable	متغير عشوائي
Unimodal	وحيد المنوال	Discrete random variable	متغير عشوائي منقطع
Harmonic mean	وسط توافقي	Dependent variable	متغير مرتبط
Median	وسط	Confidence interval	مجال ثقة
		Statistical population	مجمع إحصائي
		Set	مجموعة
Converge	يتقارب	Null set	مجموعة خالية

sampling	التوزيع العيني للنسبة	two-sided test	اختبار ثنائي الجانب
distribution of proportion		U	
sample space	فضاء العينة	unbiased estimator	تقدير غير متحيز
sampling with replacement	السحب مع إعادة	uncorrelated	غير مرتبط
sampling without replacement	السحب بدون إعادة	uncountable	غير قابل للعد
selection	اختيار	unimodal	وحيد المتوال
set	مجموعة - فئة	union	اجتماع - اتحاد
sequence	متتالية	universe	المجموعة الكلية
subset	مجموعة جزئية - فئة جزئية	V	
slope	ميل	value	قيمة
standard	معياري	variable	متغير
standard deviation	الاغراف المعياري	variance	تباين
standardized	متغير عشوائي معياري	variation	الاختلاف
random variable		vector	شعاع
stationary	مستقر	Venn diagram	مخطط فين
statistics	إحصاء	volume relatives	مناسب الحجم
statistical decisions	القرارات الإحصائية	W	
statistical	الفرضيات الإحصائية	weight	وزن
hypotheses		X	
stochastic process	عملية عشوائية	x-co-ordinate	الاحداثي السيني
step function	دالة درجية	Y	
T		y-coordinate	الاحداثي الصادي
tail	الذييل - وجه الكتابة في قطعة النقود	Z	
test	اختبار	z-co-ordinate	الاحداثي العيني
test of hypotheses	اختبار الفرضيات	zero	معاملات الارتباط من الرتبة صفر
test of significance	اختبار المعنوية	order correlation coefficients	
test statistic	إحصاء الاختبار	zero point	نقطة الصفر
theory of decisions	نظرية القرارات	zero vector	شعاع صفرى
transformed	محول		
two-tailed test	اختبار ثنائي الذييل		



measurements	قياسات	permutations	تبادل
median	الوسيط	population	مجمع إحصائي
member	عضو	probability	الاحتمال
moment	العزم	probability density function	دالة الكثافة الاحتمالية
moment about the mean	العزم حول الوسط	probability distribution	التوزيع الاحتمالي
moment about the zero	العزم حول الصفر	probability function	دالة احتمالية
multinomial	كثير حدود	product	جداء - ضرب
mutually exclusive events	حوادث متنافية تبادلها	power	قوة

## N

negative	سالب
negative distribution	التوزيع السالب
nondiscrete	غير متقطع
normal distribution	التوزيع الطبيعي
null hypotheses	فرض العدم
null set	المجموعة الخالية
number of degrees of freedom	عدد درجات الحرية

## O

object	شيء - هدف
observation	ملاحظة
one-tailed test	اختبار وحيد الذيل
one-sided test	اختبار وحيد الجانب
operation	عملية
ordered partitions	تجزئات مرتبة
ordinate	الأحداثي الصادي
outcome	نتيجة

## P

partition	تجزئة
perfect	تام

## Q

quadratic mean	الوسط التربيعي
----------------	----------------

## R

random variable	متغير عشوائي
random sample	عينة عشوائية
range	مدى
real number	عدد حقيقي
rectangle	مستطيل
rectangular-coordinates	الإحداثيات المتعامدة

region	منطقة
region of rejection of the hypothesis	منطقة رفض الفرضية

regression	الترجيح
relative frequency	التكرار النسبي
repeated tests	اختبارات متتالية
repetition	تكرار

## S

sample	عينة
sampling distribution	توزيع المعاينة
sampling distribution of mean	التوزيع العيني للوسط

density function	دالة الكثافة	<b>H</b>	
dependent variable	متغير مرتبط	harmonic mean	الوسط التوافقي
descriptive statistics	الإحصاء الوصفي	head	الرأس (وجه النقش في قطعة النقود)
diagram	شكل بياني	hypothesis	فرضيات
difference	فرق	hypergeometric distribution	التوزيع الهندسي الزائدي
discrete random variable	متغير عشوائي منقطع		
dispersion	الانتشار	<b>I</b>	
distribution	التوزيع	identity	وحدة
		impossible event	حدث مستحيل
<b>E</b>		independent	مستقل
elementary event	حدث أول	independent events	حوادث مستقلة
empirically distribution	التوزيع التجريبي	independent variables	متغيرات مستقلة
estimation	تقدير	inequality	متباينة
even	زوجي	infinite	غير محدود (لا نهائي)
event	حدث	initial	أول
expectation	التوقع	intersection	تقاطع
expected value	القيمة المتوقعة	item	عنصر — بند — فقرة
experimental	توزيع المعاينة التجريبي		
sampling distribution		<b>J</b>	
		joint probability distribution	التوزيع الاحتمالي المشترك
<b>F</b>			
factorial	عامل		
finite	محدود	<b>L</b>	
frequency	تكراري	line graph	خط بياني
frequency function	دالة تكرارية	linear function	دالة خطية
		list	قائمة
<b>G</b>			
games of chance	ألعاب الحظ	<b>M</b>	
gamma distribution	توزيع غاما	main	رئيسي
Gaussian distribution	توزيع غوص	main diagonal	القطر الرئيسي
geometrical measurement	مقياس هندسي	marginal distribution	التوزيع الهامشي
graph	شكل بياني	matrix	مصفوفة

## إنجليزي/عربي

<b>A</b>		chi-square distribution	توزيع كاي - مربع
abscissa	الاحداثى السينى	coefficient	معامل
absolute value	القيمة المطلقة	closed interval	مجال مغلق
algebra	الجبر	column	عمود
algebraic function	دالة جبرية	component	عنصر
alternative hypothesis	الفرض البديل	conditional distribution	التوزيع الشرطى
angle	زاوية	conditional probability	الاحتمال الشرطى
analytic geometry	هندسة تحليلية	confidence interval	مجال ثقة
analysis	تحليل	confidence level	مستوى ثقة
angular velocity	السرعة الزاوية	constant	ثابت
approximation	تقريب	continuous	مستمر
asymtote	خط مقارب	continuous function	دالة مستمرة
<b>B</b>		continuous probability distribution	التوزيع الاحتمال المستمر
biased estimator	تقدير متحيز	continuous variable	متغير مستمر
base	أساس	converge	يتقارب
belong	ينتمى	countable	قابل للعد
bending	انحناء	count	يعد - يحصى
Bernoulli distribution	توزيع برنولى	counting	العد - الترقيم
binomial	حدائى	critical region	منطقة حرجة
binomial coefficients	معاملات حدائى	critical value	قيمة حرجة
binomial distribution	توزيع حدائى	cumulative distribution	توزيع تراكمى
binomial expansion	مفكوك حدائى	<b>D</b>	
<b>C</b>		deductive statistics	الإحصاء الاستنتاجى
central limit theorem	نظرية النهاية المركزية	defective	معيب
chain	سلسلة		
characteristic function	دالة مميزة		









Bibliotheca Alexandrina



0450909

مطابع جامعة الملك عبد العزيز